Inequality of Opportunity: Empirical Analyses of Nepal

機会の不平等:
ネパールを事例とした実証分析

SEBASTIAN PETER SCHOLL
セバスチャン　ペータ　ショル

13 December 2019

4014S007-3
Doctor of Philosophy Program
Major in International Studies

GRADUATE SCHOOL OF ASIA-PACIFIC STUDIES
WASEDA UNIVERSITY
Tokyo, Japan

C.E. Prof. Shujiro Urata

D.E. Assoc. Prof. Kaoru Nabeshima
Prof. Atsushi Kato
Prof. Haruko Noguchi

DISSERTATION SUMMARY

Keywords: Absolute/Relative indices of inequality of opportunity, inequality decomposition, inequality in outcome

Inequality in opportunities is the core thought behind development economics. Most topics in development economics refer to the key concept of opportunities and access to them; such as credit market access, the inequality-poverty-economic growth triangle or child labour. Yet, the term inequality of opportunity (IOp) only became a reference with Roemer's 1998 benchmark book establishing the field of opportunity economics. Other concepts, such as intergenerational (e.g. income or educational) mobility, meritocracy, or just inequality, could be perceived as closely related but distinct from the concept of IOp. "Opportunity economists" (e.g. Bourguignon et al., 2007; Checci and Peragine, 2010) argue that these terminologies are partially incoherent because they include both static and dynamic elements, as well as inequalities in outcome and opportunity that arise due to individual choice and exogenous factors.

The literature on inequality in outcome (e.g. income) is abundant. It may, however, be argued that the distribution of particular outcomes is not appropriate for assessing the fairness of a social system or an allocation (e.g. Sen, 1985; Arneson, 1989). Inequality that arises due to differences in choice (attributes for which individuals can be held responsible and accountable) and therefore different outcomes, is more ethically acceptable. In other words, income gaps that arise from differences in efforts are generally less objectionable than those that are due to differences in race or gender, for example. The goal of the decomposition of IOp is to single out to what extent inequality is due to unequal *opportunities*.

This study focuses on Nepal and primarily utilises the 2011 Nepal Living Standards Survey Round Three (NLSS III) household dataset. In order to give a comprehensive overview of Nepal and allow for adequate results interpretation, essential country background information (e.g. on history, social structure, and geomorphology) is given. The utilised framework in this study is partially based on the model developed by Ferreira and Gignoux (2011) and the theoretical concept of IOp draws on Roemer (1998). Roemer explains that differences in *outcome* or *advantages* (y) can be explained by firstly, exogenous *circumstances* variables (C) to an individual such as race and secondly, an endogenous *effort* variable (e) for which an individual can be held responsible and accountable. Perfect equality of opportunity exists when outcomes are distributed independently of circumstances and solely depend on efforts.

The main analysis of this dissertation has two dependent outcome variables (y): per capita consumption expenditure and per capita income, and seven circumstance variables: gender, ethnicity, religion, father and mother's education, and two birth region groupings. Individuals are divided into different *types* or *groups* in which they are homogenous in terms of theirs circumstance characteristics. Individual i is characterised by (y_i, C_i^j, e_i). The empirical analysis mainly draws on two models used by Ferreira and Gignoux (2011). Firstly, a model inspired by Bourguignon et al. (2007), a linear regression and inequality decomposition model based on the following reduced form: $y=f[C,E(C,v),u]$. Secondly, a model inspired by both Bourguignon et al. (2007) and Checci and Peragine (2010), a parametric and non-parametric approach for both absolute (IOL) and relative (IOR) IOp indices. IOR measures the level of IOL in relation to total inequality. These indices exclude the endogenous effort variable and satisfy the standard axiomatic properties in the literature on the measurement of relative inequality. IOR, IOL, various income disparity measures (e.g. Gini coefficient, Mean Log Deviation) and the OLS regression estimates are compared. IOp indices are computed for the national and the regional development level, urban and rural areas, and for economic outcome groupings by population quarters. The results are checked for robustness by using equivalence scale and other measures.

	Parametric model	Non-parametric Model
IOL	$\theta_a^{PS} = E_0(y) - E_0(\tilde{v})$	$\theta_a = E_0(\{\mu_i^k\}) - \frac{1}{N}\sum_{i=1}^{N} \log \frac{\mu}{\mu_i^k}$
IOR	$\theta_r^{PS} = 1 - \frac{E_0(\tilde{v})}{E_0(y)}$	$\theta_r = \frac{E_0(\{\mu_j^k\})}{E_0(y)}$

μ is the mean group advantage for k type and l type, E_0 indicates the inequality index in advantage y, and the smoothened and standardised advantage v.

These pragmatic and relatively simple scalar measures of IOp also allow for analysing the most opportunity-deprived types and their specific circumstantial profiles for economic outcome. The cross-comparison of these results is of direct relevance for formulating customised and localised opportunity-equalising policies.

The results suggest that the level of IOp at the national level is lower than most of those found in the literature, with opportunity deprivation accounting for 26% to 32% of overall inequality in outcome. The geographically disaggregated results, however, indicate a large divergence across the country. Urban areas, for instance, face opportunity deprivation levels reaching between 34% to 44% out of total outcome inequality (vs. 17% to 22% in rural areas). Interestingly, the circumstantial decomposition suggests that negatively impacting factors (e.g. lower caste group belonging or absence of parental education) have a) a more ample effect in urban centres, and this is b) disproportionately aggravated by the accumulation of such negative factors. Also, the top 25% of the population by consumption faces a higher level of within-group IOp than the poorest 25% of the country.

The circumstance-specific analyses suggest that geographical disaggregation is crucial. There are some counterintuitive results. 1) Gender is consistently insignificant across all analyses. 2) Father's education seems more important on impacting on outcome than mother's education. 3) Buddhists, despite being a religious minority, are afforded an elite and opportunity-enhancing status. 4) Caste group belonging still largely predicts the access to opportunities and this is more so in urban than rural areas, despite Nepal having abandoned this societal stratification system decades ago. 5) At the national level, being born in the countryside negatively affects economic outcome, but when disaggregated, the most disadvantaged types suffer less in rural areas. In general, the most opportunity-deprived types are individuals who accumulate four or more of the opportunity-depriving circumstances, but the most disadvantaged types suffer disproportionately more in the central region and in urban areas. This suggests that either they should stay in the countryside or emigrate.

The results suggest that the country may be trapped, and all parts of society suffer. It appears that geographical mobility acts as an opportunity equaliser, no matter the socio-economic status. Both, the disadvantaged and the most advantaged either migrate or emigrate. Yet it may be the existence of these inequalities that trap the country at a low stage of development, with over 25% of the population emigrating to try to overcome opportunity deprivation at home and the government failing to combat the depopulation of the young and the issue of brain drain.

Moreover, when this study's results are compared to the literature, they suggest that Nepal may face an increase in both inequality of opportunity and outcome as the country continues to develop. An adaptation of the Kuznets curve for IOp is also proposed.

Main references

Arneson, R. (1989), "Equality and Equal-Opportunity for Welfare". Philosophical Studies, 56(1), 77-93.

Asadullah, M.N. and Yalonetzky, G. (2012). "Inequality of Educational Opportunity in India: Changes Over Time and Across States", World Development, 40(6), 1151-1163.

Bourguignon, F. Ferreira, F.H.G. and Menendez, M. (2007). "Inequality of Opportunity in Brazil", Review of Income and Wealth 53(4), 585-618.

Checci, W., Peragine, V. (2010). "Inequality of Opportunity in Italy", Journal of Economic Inequality 8(4), 429-450.

Ferreira, F.H.G. and Gignoux, J. (2011). "The Measurement of Inequality of Opportunity: Theory and an Application to Latin America", Review of Income and Wealth, 57(4), pp. 622-657.

Roemer, J. E. (1998), "Equality of Opportunity", Harvard University Press.

Sen, A. (1985), "Commodities and Capabilities", North Holland, Amsterdam.

ACKNOWLEDGEMENTS

"Our lives are shaped by opportunities. How we perceive them. If we dare to seize them.
I was excited to enrol into the oldest Italian university after completing high school.
The first to study abroad in my family, and that as a woman.
But destiny took its turn…
Now, you are able to fulfil this dream for me."

– My great-grand mother at the age of 98, just weeks before her passing. (referring to 1939, Germany)

My numerous travels to Latin America and Africa were the source of inspiration to study the concept of inequality of opportunity. Some of the most intelligent people I have met (on a human, moral and intellectual level) come from these regions. Yet something stops them from progressing, from moving on. They have little prospects, little opportunities. If they had the same possibilities as me (coming from a developed country) they would thrive.

Gender economists have long argued that not including women into the work force is economically irrational. It is not just women, but also other groups that are excluded, for instance, from access to education, health care or the credit market. Not only are they unable to contribute to the fullest of their capacities to the overall aggregate economic performance, but they are also unable to seek complete personal fulfilment. So why just fight for gender equality, when numerous exogenous other factors also lead to opportunity deprivation. Hindering individuals and groups from fulfilling their own dreams is of loss to them and to the nations they live in. I want to demonstrate that not seizing the incredible economic potential of the largely young and ambitious people is unwise. If it were possible to persuade policy makers of the capabilities and potential of their own people, and how they can contribute to their own success, development would turn into an unstoppable avalanche of equitable and inclusive economic prosperity.

Upfront I would like to thank Professor Fuwa for believing in me and without whom I would have never even considered doing a PhD. His academic integrity, honesty, dedication towards his students and respect will always be an inspiration to me. A grand thank you also goes to Professor Urata and Professor Nabeshima for their valuable and dedicated advice and clear guidance over the years. Also, I want to express my gratitude to Professor Noguchi and Professor Kato for their comments and suggestions on my research and development progress while having figured on my PhD guidance committee.

I also want to thank Waseda University and its outstanding GSAPS staff who have made my time there unforgettable. Without their continuous support and the institutional framework I would have not been able to broaden my horizons to the current extent. I would have never had the chance to fully indulge in Asian regional integration without the support of the East Asia University Institute for Asian Regional Integration Program and Peking University, where I was welcomed with open arms for one semester exchange.

Waseda has also helped me in finding financial and academic support. The Waseda Scholarship for Young Doctoral Students and the JASSO Honors Scholarship guaranteed for a smooth research environment. The Haraguchi Memorial Asia Research Fund allowed me to dig deep into the development challenges on the ground during my field research, without which my research would not have been complete. The Tonen International Scholarship Foundation (**公益財団法人東燃国際奨学財団**) did not just provide financial support, but the institutions encouragement through its regular meetings and study trips further helped me to maximise my time as a graduate student in Japan, for which I am truly grateful.

I would like to express my sincerest gratitude to all institutions, officials, academics, entrepreneurs and individuals who frankly opened up to me while I was doing my field research in Nepal. A particular thank you goes to Narayan Poudel and his family, and the officials at Nepal's Central Bureau of Statistics and at the National Planning Commission.

I want to thank those who have opened my eyes, heart and thoughts. A special thanks goes to my father and mother, who have always given me the freedom to take my own decisions and taught me that not only do decisions have to be taken independently, but that one must hold responsibility for one's own actions or inactions. My friends and particularly my girlfriend, CV, who has always been by my side and I can count on anytime. Our mutual encouragement and inspiration are priceless.

Last but not least, I would like to thank my Zemi friends (alumni and current), who were of great academic and moral support over the years. They all come from diverse backgrounds, hold various professional positions, and have all inspired me with their motivations to drive development. I will always cherish these fostered friendship. In particular, I want to thank Harue, Adam, Pipe and Yeng-Tsung. Words don't do their support justice.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

CBS	Central Bureau of Statistics (Government of Nepal)
DU	Direct Unfairness Approach (Romer and Trannoy, 2015)
EOp	Equality of opportunity
EOP	Equality of opportunity policy
FDI	Foreign Direct Investment
FG (2011)	Ferreira and Gignoux (2011)
GDP	Growth Domestic Product
GE	Generalised Entropy
HH	Household
HHH	Household head
HCI	The World Bank's *Human Capital Index*
HDI	Human Development Index
HOI	The World Bank's *Human Opportunity Index*
IGM	Intergenerational mobility
IOL	Absolute level of inequality of opportunity
IOp	Inequality of opportunity
IOR	Relative inequality of opportunity
LSMS	Living Standard Measurement Surveys
MDGs	Millennium Development Goals
MLD	Mean Logarithmic Deviation
NLSS I	Nepal Living Standard Survey I (1995/1996)
NLSS II	Nepal Living Standard Survey II (2004/2005)
NLSS III	Nepal Living Standard Survey III (2010/2011)
NRs.	Nepalese Rupees
OLS	Ordinary Least Square
RF	Reduced Form
SDGs	Sustainable Development Goals
SEZ	Special Economic Zones
USD	United States Dollar
WB	World Bank
WDI	World Bank Development Indicators

Chapter 1 Introduction

This chapter introduces the study of inequality in outcome and inequality of opportunities in Nepal. It first gives a brief insight into the study. It then outlines the research questions, hypotheses, literature contributions, and findings for the inequality in outcome analyses. The same is next done for the study of inequality of opportunity. Finally, the research methodology and dissertation structure are given.

Inequality in opportunities is the core thought behind development economics. Most topics in development economics refer to the key concept of opportunities and access to them; such as credit market access, the inequality-poverty-economic growth triangle, or child labour. Yet, the term inequality of opportunity (IOp) only became a reference with Roemer's 1998 benchmark book establishing the field of opportunity economics. Other concepts, such as intergenerational (e.g. income or educational) mobility, meritocracy, or just inequality, could be perceived as closely related but distinct from the concept of IOp. "Opportunity economists" (e.g. Bourguignon et al., 2007; Checci and Peragine, 2010; Chetty et al., 2014) argue that these terminologies are partially incoherent because they include both static and dynamic elements, as well as inequalities in outcome and opportunity that arise due to individual choice and exogenous factors.

The literature on inequality in outcome (e.g. income) is abundant. It may, however, be argued that the distribution of particular outcomes is not appropriate for assessing the fairness of a social system or an allocation (e.g. Sen, 1985; Arneson, 1989). Inequality that arises due to differences in choice (attributes for which individuals can be held responsible and accountable) and therefore different outcomes, is more socially and ethically acceptable. In other words, income gaps that arise from differences in efforts are generally less objectionable than those that are due to differences in race, parental background, or gender, for example. The goal of the decomposition of IOp is to single out how much inequality is due to unequal *opportunity*. Thus, the opportunity indices in this study try to present an alternative and more comprehensive measure of inequality, more adapted to the contemporary ideals and norms of the theory of distributive justice.

Despite the gradual shift in the literature of distributive justice away from the traditional analysis of inequalities in economic outcome to inequalities in economic inputs, the academic literature on IOp is relatively recent and scarce at the theoretical level and even more so on the empirical level. Most of the scarce empirical literature focuses on developed countries, such as the US and some European economies. This study is primarily an empirical extension to a developing country, Nepal.

The opportunity indices of this study are computed at the national level but are also disaggregated geographically and by their input factors. In total, there are four separate levels of econometric analyses that study the existence, amplitude, and dispersion (i.e. concentration) of IOp in Nepal: the national level, Development Region level, urban rural area level, and population grouping by economic welfare quarter level analysis.

These, to some extent, cross-country comparable indices, as well as their disaggregation, allow for the elaboration of more precise and custom-tailored policies and projects to combat inequalities in opportunities and guarantee inclusive development. Consequently, these should enhance living standards, especially for the most impoverished and discriminated. Opportunity-equalising policies should focus on and encourage all individuals to engage in economic activities to the best of their abilities, in return boosting overall national economic performance, social inclusion, and political stability, and enabling personal fulfillment.

1.1. Research Motivation

My numerous travels to Latin America and Africa were the source of inspiration to study the concept of IOp. During my travels, it always struck me that despite some people being very capable (soft and hard skills), unequal opportunities blocked their career prospects, their ability to make a fair living and to find fulfilment.

Gender economists have long argued that not including women in the work force is economically irrational. But it is not just women but also other groups that are excluded, for instance, from access to education, health care, or the credit market. Not only are they unable to contribute to the fullest of their capacities to overall aggregate economic performance but they are also unable to seek complete personal fulfilment. So why just fight for gender equality when numerous other exogenous factors also lead to opportunity deprivation. Hindering individuals and groups from fulfilling their own dreams is of loss to them and to the nations they live in. I want to demonstrate that not seizing the incredible economic potential of the mostly young and ambitious people is unwise. If it were possible to persuade policy makers of the capabilities and potential of their own people, and how they can contribute to their own success, development would turn into an unstoppable avalanche of equitable and inclusive economic prosperity.

It is this personal curiosity that set off my interest in opportunity economics. As gender economists argue that gender inequality brings about tremendous socio-economic loss, I argue that inequality of opportunity at a more aggregate level does too.

Reading the literature on inequality, it became evident that i) ample research exists on the study of inequality in outcome; however, ii) little literature exists on IOp. Despite the benchmark concept having been coined by Roemer (1998) more than 20 years ago, Pikkety's (2014) book entitled "Capital in the Twenty-First Century" having re-heated public debates on inequality and its underlying norms in 2014, and having seen a shift in public opinions on the definition of inequalities (e.g. Germany's government placing "Chancengleichheit" (i.e. Equality of Opportunities) into the government's coalition agreements since the early 2000s), most of the literature that does exist is mainly theoretical. While Roemer's IOp concept is theoretically appealing, its operationalisation for empirical research, however, is challenging. Furthermore, Kanbur and Wagstaff (2014) argue that it may not be the most suitable analytical tool to instruct equality policies due to the lack in consensus on its underlying philosophical, ideological, and normative thoughts. Yet, it can be argued that these fundamental bases of, for instance, to which degree and level opportunities should be equalised (e.g. in access to education or healthcare), are to be decided by the people of each nation. The theory of distributive justice (e.g. Kincaid and Ross, 2009; Roemer, 1993) should serve as reference. There is a need to adapt the policies of today to the public opinions of the people.

The goal of the opportunity indices is to try to measure overall aggregate opportunity deprivation, which can then be disaggregated in terms of the circumstantial factors and help to isolate the opportunity deprivation profiles of groups of people. These results can then serve to instruct policy makers to effectively implement well-targeted policies in order to promote long-term, inclusive economic development and social stability. The degree of opportunity deprivation and the types of people who benefit or suffer the most from it are studied using an unusually rich microlevel dataset for a developing country. The study is a cross-section analysis for the Nepalese society in 2011.

There is an important gap in the IOp literature at both the theoretical and empirical levels, not just in trying to quantify opportunity deprivation and the profiles of people who suffer the most but also in trying to understand its impacts on, for example, economic growth, entrepreneurship, innovation, political or social stability, or psychological well-being.

There are various reasons as to why this study focuses on Nepal. First , this study tries to narrow the gap in the empirical IOp literature by analysing Nepal, one of the world's 30 poorest countries and a caste society. The empirical application and geographical extent of the concept of IOp is fairly sparse compared to research on standard inequality decomposition or outcome inequality for example. Its geographical coverage mainly focuses on Latin America and developed countries (e.g. Germany or US). Hassine (2011) has so far provided the only study for the Arab region. In Asia, most (of the few) IOp studies focus on India, particularly due to its caste society system, fast economic development and its relatively extensive data availability (e.g. Asahulla and Yalonetsky, 2012). The availability of extensive HH data in emerging countries is limited and the need for exhaustive data on exogenous circumstance variables and the responsibility variable poses a particularly strong barrier for empirical analysis.

Second, the analysed dataset (Nepal Living Standards Survey III; from now on NLSS III) is an unusually rich microlevel dataset in terms of information and sample size for a developing country. It contains detailed information on household roasters, individual's and households' characteristics such as education and parental background. These details are crucial for a thorough IOp analysis and are rare to obtain, especially for developing countries. The richness of the data allows to further disaggregate the national level analysis by geographical and demographic factors, which has not been done in such a thorough manner before in the literature. This in turn allows for a much more customised, localised and effective IOp policy framing. Furthermore, the richness of the data allows to relax a series of household roaster constraints (such as only being able to analyse household heads and spouses) to all household members that satisfy certain criteria later explained. This in turn renders the analysis more robust by significantly increasing the population sample and including a larger variety of household members. Many IOp papers are only able to use household heads and spouses due to the nature of their datasets (e.g. labour force surveys; FG, 2011).

Third, the IOp analysis of Nepal is an important empirical contribution to the literature of caste and ethnic group belonging in economic development. To the knowledge of the author very few studies that link IOp and caste and ethnic group belonging exist (e.g. Asahulla and Yalonetsky, 2012).

1.2. Research Questions, Hypotheses, Main Findings, and Literature Contributions

1.2.1. Inequality in Outcome for Income and Consumption Expenditure: Analyses of Nepal

This section focuses on the research questions, hypotheses, findings, and main literature contributions of the inequality in outcome chapter.

1.2.1.1. Inequality in Outcome: Research Questions

There are three main clusters of research questions to which the inequality in outcome chapter attempts to find answer, and close the gaps in the corresponding literature.

Q1: Which welfare measures are the most adequate for capturing the economic livelihood of the Nepalese people? And how are these calculated?

Q2: How is income and consumption expenditure distributed across Nepal? Which parts of the population earn or consume how much?

Q3: Which inequality in outcome index is the most suitable to serve as a base reference for the subsequent inequality of opportunity analyses?

1.2.1.2. Inequality in Outcome: Research Hypotheses

Chapter 2 of this study examines whether there is inequality in outcome (i.e. inequality of per capita income or inequality of per capita consumption expenditure) by testing the null hypothesis that the Mean Logarithmic Deviation (MLD) is equal to zero (H_0: MLD = 0). Under the null, there is perfect outcome equality. Note that the MLD of the Generalized Entropy inequality indices is used as the main reference inequality in outcome measure (i.e. index) for the subsequent IOp analyses in this study.

Hypotheses:

H_0: MLD = 0, when MLD = 0, then there is perfect equality of outcome.
H_1: MLD > 0, when MLD > 0, then the null is rejected and inequality in outcome exists.

1.2.1.3. Inequality in Outcome: Main Findings and Literature Contributions

Table 1.1 summarises the main findings and literature contributions of the chapter on inequality in outcome. A detailed elaboration can be obtained in Chapter 2 and Chapter 4: Inequality in Outcome and Inequality of Opportunity: Results Comparison and Literature Contributions.

The literature on inequality in economic outcome, namely income and consumption inequality, is ample, both on a theoretical and an empirical level. The study of inequality is at the core of economic development. Much of it is considered as acquis, however, it was to the author's surprise to uncover that there is little consensus on the aggregation methodology used to aggregate the variables that then serves as key reference in calculating the inequality indicators. In other words, there is little consensus in the literature on how to aggregate, for instance, income and consumption expenditure. This study digs into a) the literature on the measure of (economic) welfare to justify the choice of income and consumption expenditure as economic welfare proxies. Most of the literature does not debate as to why these two economic welfare measures are used in their analyses, while this study places significant emphasis on the arguments for and against them.

b) This study undertakes a thorough literature review and comparison of the most prominent methodological guidelines on the aggregation of income and consumption in order to optimise the approach for the NLSS III dataset and the Nepal-specific context. None of the studies the author came across discuss their aggregation methodologies for the economic welfare measures, despite them also using microlevel data and the important effect the aggregation methodology can have on the empirical analyses.

c) An adjusted aggregation methodology is proposed for both income and consumption. These approaches take into consideration i) the particularities of the NLSS III dataset, and ii) the Nepal-specific context. Such elements include adjustments to regional living standards, inclusion or exclusion of health and educational expenses, and unit measurement conversions.

d) Variations of the income and consumption aggregates are calculated, following not just the recommendations in the literature but the author's fine-tuned approaches. These are checked for

robustness and consistency at every level of the aggregation process (e.g. unit measurement checks, bootstrapped standard errors for indices, outlier checks, adjusting for different costs of living). From the empirical results, a stark variation in the aggregates is revealed depending on the aggregation methodology followed. This underpins the validity of the chosen adjusted methodologies for the author's recommend consumption and income reference variables.

e) The careful choice of the aggregation methodology and the consequent outcome measures, i.e. the chosen income and consumption aggregates, is essential for the following inequality in outcome and the inequality of opportunity analyses. Previous IOp studies do not appear to place emphasis on this, and just take the aggregate levels of income and/or consumption for granted without greatly discussing them. However, this has a significant impact on the further inequality (in outcome and opportunity) analyses. The impacts are highlighted by empirically showing the impacts the aggregations methodologies also have on i) the amplitude of the chosen inequality measures, ii) the divergence between the inequality measures, and iii) the divergence between the inequality indices depending on whether consumption or income serve as the main economic welfare measures.

f) A thorough literature review of the most commonly is adhered to and referenced inequality in outcome measures are done. The most suitable inequality in outcome indicator is isolated from both a theoretical and an empirical level. The chosen inequality in outcome index is the Mean Logarithmic Deviation (MLD). This is a commonly referred to index in the inequality in outcome literature. It also serves as the inequality in outcome reference index for FG (2011). However, they do not explain or greatly discuss the extent of the use of the MLD. This study contributes in giving a more elaborate debate on the trade-offs of using the MLD. This is important since the most intuitive and relevant IOp results (particularly from the Inequality of Opportunity Ratio (IOR)) are obtained and interpreted by using the used inequality in outcome (i.e. the MLD) as reference. There are two of the key advantages of the ML. First, the MLD is the only measure of the Generalised Entropy (GE) indices of inequality to satisfy the path independence axiom. A more detailed elaboration on this later. Second, it is also the only measure of the Generalised Entropy (GE) indices that is decomposable between income sources and between individuals, groups and sub-groups. It is the latter that is of particular interest for the IOp study, which divides the population into "types" or groups of individuals in which they are homogenous in terms of circumstances. A full elaboration on this can be found in Section 3.1, which gives a detailed literature review of IOp, and Section 3.2, which describes and explains the IOp methodology applied to the NLSS III dataset in this dissertation.

Table 1.1: Inequality in Outcome: Summary of Main Findings and Literature Contributions

Literature	Contributions
Inequality in Outcome Methodology and Empirics	
Ample literature on income and consumption inequality (theory and empirics), but little consensus on how to aggregate economic outcome, which in turn influences the inequality indices' computation.	- Literature review on the measure of (economic) welfare, and justification for the utilization of income and consumption as the most adequate economic welfare measures. - Literature and guideline review of the income and consumption expenditure aggregation methodologies - Adjusted aggregation methodology developed for the NLSS III dataset and the Nepal-specific context → important for inequality in outcome indices → important for subsequent IOp analysis - Thorough computation and numerous robustness checks (e.g. unit measurement checks, bootstrapped standard errors for indices, outlier checks, and adjusting for different costs of living) of the economic outcome aggregates and aggregate results comparison. - Thorough computation and results comparison of various inequality in outcome indices. - Theoretical and empirical discussion on the use of the MLD as the reference inequality in outcome variable for the IOp analyses

1.2.2. Inequality of Opportunity in Nepal: National Level, Development Region Level, Urban vs. Rural Areas, Income and Consumption Expenditure Groupings by Quarters

This section focuses on the research questions, hypotheses, findings, and main literature contributions of the IOp chapter.

1.2.2.1. Inequality of Opportunity: Research Questions

Broadly speaking, there are four main research questions concerning the study of IOp in Nepal. The goal is to answer them as precisely as possible and close the corresponding gaps in the literature:

Q1: Does IOp exist in Nepal?
Q2: If IOp does exist, what is its amplitude?
Q3: How is IOp distributed / concentrated across the country?
Q4: Who are the groups of people who are the most discriminated against?

In an attempt to reply to the research questions as precisely as possible, the IOp chapter has four different levels of analyses, each containing rigorous descriptive and econometric elements. The four levels are:

1) National level analyses
2) Development Region level analyses
3) Urban rural area level analyses
4) Population grouping by economic welfare quarter analyses

There are two main clusters of research questions for each of these levels.

1) National level analyses

   Q1: Does IOp exist? And if so, what is its amplitude?
   Q2: Which groups of people are the most opportunity deprived at the national level? What are their relative mean economic outcomes?

2) Development Region level analyses

   Q1: Does IOp exist? And if so, are there any differences in amplitudes between the five Development Regions of Nepal? Which region has the highest degree of IOp and which has the lowest?
   Q2: Which groups of people within each Development Region are the most opportunity deprived? What are their relative mean economic outcomes?

3) Urban rural area level analyses

   Q1: Does IOp exist? And if so, is there any difference in the amplitudes between the urban and rural regions? Which area faces the highest degree of IOp?
   Q2: Which groups of people within the urban and rural areas are the most opportunity deprived? What are their relative mean economic outcomes?

4) Population grouping by income and consumption quarter analyses

   Q1: Does IOp exist? And if so, are there any difference in amplitudes between the earner and consumer groups? Which quarter of the population faces the highest degree of within-group IOp and which the lowest?
   Q2: Which groups of people within each economic outcome quarter are the most opportunity deprived? What are their relative mean economic outcomes?

1.2.2.2. Inequality of Opportunity: Research Hypotheses

This study examines whether there is IOp in Nepal by testing the null hypothesis that the scalar indices of inequality of opportunity (IOL and IOR) are equal to zero (H_0: IOL = 0 and H_0: IOR = 0). Under the null, there is perfect equality of opportunity, and circumstances do not affect economic outcome (i.e. per capita income or per capita consumption expenditure).

Hypotheses to test for Q1s above i.e. "Does IOp exist? And if so, is there any difference in the amplitude between […]":

H_0: IOL = 0 and **H_0: IOR = 0**, when IOL or IOR=0, then there is perfect equality of opportunity.
H_1: IOL > 0 and **H_1: IOR > 0**, when IOL or IOR>0, then the null is rejected and IOp exists.

In case the null hypothesis is rejected (IOL or IOR > 0), the question is to what extent each circumstance variable (C) affects economic outcome. This is done by running the following RF-OLS regression:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + u_i$$

Where

y_i	Per capita consumption expenditure or per capita income of the individual i.		
β_0	Constant	X_{5i}	Father primary incomplete or above
β_j	Circumstance variable parameter (j = 1, … 9)	X_{6i}	Mother primary incomplete or above
X_{1i}	Female	X_{7i}	Birth region rural
X_{2i}	Ethnic minority/dominated caste group	X_{8i}	Birth region hill
X_{3i}	Buddhist	X_{9i}	Birth region Terai
X_{4i}	Religion other	u_i	Error term

B_j represents the parameter for each exogenous categorical circumstance variable X_j ,and *i* indices the individuals under study:

H_0: $\beta_j = 0$, Circumstance *j* does not impact on economic outcome.
H_1: $\beta_j \neq 0$, Circumstance *j* impacts on economic outcome and the null is rejected.

The goal of this study is to identify which types or groups *k* of individuals are the most opportunity-deprived, based on the estimated regression coefficients of the circumstance variables that are statistically significantly different from zero ($\beta_j \neq 0$).

Testing for the partial IORs, that is the circumstantial specific share of overall outcome inequality (i.e. IOR).

H_0: partial $IOR_j = 0$, when **partial $IOR_j = 0$,** then the circumstance specific share of circumstance *j* of overall outcome inequality is null.

H_1: partial $IOR_j \neq 0$, when **partial IOR ≠ 0**, then the null is rejected and the circumstance specific share of circumstance *j* of overall inequality in outcome is different to zero, i.e. the circumstance explains a certain share of overall inequality in outcome, suggesting the importance of the circumstantial categorical variable for the accessibility in opportunities.

The RF-OLS regression and the partial IORs add additional depth of analysis to the computed scalar indices and allow to identify social groups at which equality of opportunity policies should be aimed. Roemer (1998) refers to such policies as *Equality of Opportunity Policies* (EOPs) as a set of allocation rules that aim to maximise economic outcome for the worst-off type(s).

Note that the above is valid for all four levels of analyses, only that the urban rural level excludes the birth region 1 (urban rural area) categorical circumstance variables.

1.2.2.3. Inequality of Opportunity: Main Findings and Literature Contributions

Table 1.2 summarises the main findings and literature contributions of the chapter on IOp. A detailed elaboration can be obtained in Chapter 4: Inequality in Outcome and Inequality of Opportunity: Results Comparison and Literature Contributions.

Most of the IOp empirical literature focuses on developed countries, such as the US or European economies. Relatively little empirical literature exists for developing countries. This is mainly due to the lack of availability of exhaustive and rich microlevel datasets. This study contributes to the literature by being an empirical extension to developing Asia, namely Nepal. Access to a relatively rich micro dataset has allowed the author to exploit it for the analysis of IOp. The benefit of this geographical empirical extension is twofold. First, it is an additional Asian country that is analysed. Second, it is a caste society.

The literature on IOp tends to compute nation-wide inequality in opportunity profiles for types of people, which are then disaggregated in terms of their circumstances or earnings, or by geographical factors. This study, however, adds three layers of analysis to the national-level IOp investigation. This is done by also studying the amplitude of IOp at two different geographical levels (1) the Development Region level and 2) the urban versus rural area level and at the population groupings level (division of the population into income and consumption quarters). In other words, there are four levels of analyses for the study of IOp in Nepal, all of which complement each other and add depth to the study. At each of the three disaggregated levels, the within and between group IOp is tested. This disaggregation is Nepal specific and cannot directly be compared to other countries or regions. Some general suggestions may, however, be derived. More on this in the results' value-added sections.

Some IOp literature focuses on specific sub-categories of opportunity deprivation, such as inequality in educational or health opportunity. This study takes another empirical approach due to the richness of the dataset, in computing overall IOp indices that try to quantify the globality of opportunity deprivation in relation to total inequality of outcome. In other words, it tries to quantify how much of total economic welfare inequality is due to opportunity deprivation. The emphasis here is very different to what the sub-categorical opportunity studies try to show.

The results suggest that the level of IOp at the national level is lower than most in the literature, with opportunity deprivation accounting for 26% to 32% of overall inequality in outcome. The geographically disaggregated results, however, indicate a large divergence across the country in terms of amplitude and concentration. Urban areas, for instance, face opportunity deprivation levels reaching between 34% to 44% out of total outcome inequality (vs. 17% to 22% in rural areas). Interestingly, the circumstantial decomposition suggests that negatively impacting factors (e.g. lower caste group belonging or absence of parental education) have, first, a more ample effect in urban centres, and this is, second, disproportionately aggravated by the accumulation of such negative factors. Moreover, the top 25% of the population by consumption faces a higher level of within-group IOp than the poorest 25% of the country.

The circumstance-specific analyses suggest that geographical disaggregation is crucial. There are some counterintuitive results. 1) Gender is consistently insignificant across all analyses. 2) Father's education seems to be more important on impacting on outcome than mother's education. 3) Buddhists, despite being a religious minority, are afforded an elite and opportunity-enhancing status. 4) Caste group belonging still largely predicts access to opportunities, and more so in the urban than rural areas, despite Nepal having abandoned this societal stratification system decades ago. 5) At the national level, being born in a rural area appears to negatively impact on economic outcome. However, when disaggregated, the disadvantaged suffer less in rural areas, and thus rural birth region is seen as positive. In general, the most opportunity-deprived types are individuals who accumulate four or more of the opportunity depriving circumstances, but the most disadvantaged types suffer disproportionately more in the central region and in urban areas. This suggests that either they should stay in the countryside or emigrate.

The results suggest that the country may be trapped, and all parts of society suffer. It appears that geographical mobility acts as an opportunity equaliser, no matter the socio-economic status. Both,

the disadvantaged and the most advantaged either migrate or emigrate. Yet it may be the existence of these inequalities that trap the country at a low stage of development, with over 25% of the population emigrating to try to overcome opportunity deprivation at home and the government failing to combat the depopulation of the young and the issue of brain drain.

The results suggest that Nepal is trapped and caught in a four-way downward spiral. 1) The most disadvantaged migrate and try to emigrate in order to make an honest living and desperately try to equalise their opportunities. However, this often entails debt and overburdening of the remaining family members, etc., which leads to further aggravation and opportunity deprivation. 2) The most advantaged also try to emigrate leading to brain drain and a challenge for the development of a healthy dynamic economy at home. 3) Nepal is facing the rapid development of a dual education system (public and private). This decreases the responsibility and willingness for public institutions to invest their limited resources, which incentivises foreign aid and, thus, further foreign aid dependence. 4) Remittances have become an essential pillar to Nepal's economic development strategy, undermining efforts to try to diversify its domestic economy and increase employability rates at home.

The findings of this study help to reduce the literature gap on IOp and provide some crucial decision-making information for policy makers. The implications here, however, are limited to the geographical extent of the study.

Table 1.2: Inequality of Opportunity: Summary of Main Findings and Literature Contributions

Literature	**Contributions**
Inequality of Opportunity Methodology	
IOp Methodology (three main strands) 1. Linear regression for economic outcome in terms of circumstances and efforts. Use of hypothetical distributions. Consequent earning decomposition. Direct and indirect effect of circumstances (e.g. Bourguignon et al., 2007) 2. Between-group inequality decomposition. Ex-ante and ex-post measures of IOp (e.g. Checci and Peragine, 2010) 3. Stochastic dominance comparison of distributions by type (e.g. Lefranc et al., 2008)	**IOp analyses** 1. Combination of IOp analyses approaches 1 to 2 and inspiration from Ferreira and Gignoux (2011) and **adaptations to the NLSS III dataset** and **Nepal-specific country context** (e.g. choice of circumstances and categorical variables). Direct ex-ante approach of the compensation principle. 2. **Parametric and non-parametric** IOp indices 3. **Additional robustness checks** (e.g. regression tests and results comparison between different level analyses) 4. **Population sample diversification and relaxation of data constraints** (e.g. not just HH heads and spouses, but all HH members)
Inequality in Outcome and Inequality of Opportunity: General Empirics	
- Most literature **focuses on the developed world** e.g. the US and EU economies. Few analyses on the developing world due to data constraints. Even WB's HOI Report mainly focuses on LAC and sub-Saharan Africa. - **Caste societies**: main focus on India (by province). - **IOp: national indices** and then within inequality decomposition.	1. **Geographic extension**: developing Asia, Nepal. Relatively exhaustive HH level dataset → Nepal-specific context. 2. **Caste society extension**. The only other IOp studies that include caste belonging are those on India. 3. **IOp level disaggregation**: National IOp analysis (as in other empirical literature) for Nepal, but further disaggregated computation by Development Region, urban vs. rural areas, and income and consumption groupings (quarters). In total: four levels of IOp analyses. Within-group and between-group analyses. 4. **Overall IOp indices** including various opportunity depriving/enhancing elements (i.e. circumstances). Further depth is added with RF-OLS regression analyses of circumstances on economic outcome and circumstance specific shares (partial

- **Focus on specific opportunities** e.g. educational, health, or labour access opportunities.	IORs). Isolation of opportunity profiles by type and by economic advantage.

Inequality in Outcome and Inequality of Opportunity: Results

Analyses Disaggregation Results

1. Important **divergences in terms of economic welfare across the country**. Variance in both level and concentration of economic welfare.
2. **Importance of geographical disaggregation** in understanding the local reality. "One size fits all policy" would be ineffective.
3. **IOp national level relatively low** (26-32%) compared to the literature → geographical disaggregation paints a different picture: important divergences by geographical region (e.g. urban (34-44%) vs. rural (17-22%)) and type of individuals.
4. **Consumption quarters:** bottom 25% of the population face a bare 3-9% vs. top population quarter 7-18% of within-group opportunity deprivation out of overall outcome inequality.

Counterintuitive Results

1. **Gender**: mostly insignificant in IOp and OLS analyses, but important gender gap for the most impoverished in the central region.
2. **Parental education**: father's education tends to be more important than mother's education
3. **Religious minority**: Buddhists are afforded an elite and opportunity-enhancing status
4. **Caste group belonging**: still matters, yet more in urban than rural areas
5. **Rural birth region**: at the national level, being born in the countryside negatively affects economic outcome, but when disaggregated, the most disadvantaged types suffer less in rural areas.

Other Interesting Results

1. In general, across all analyses, the mean economic outcome drops most severely when **four or more opportunity-depriving circumstances** are accumulated. Disproportionately negative in the central region and urban areas → should either emigrate or remain in the countryside.
2. Nepal is in a **poverty and development trap** due, in part, to opportunity deprivation → migration and remittance dependence, foreign aid dependence, social networks acting as socio-economic safety nets, etc.
3. **Geographical mobility as opportunity-equalising instrument**: both the most disadvantaged and the most advantaged migrate (domestically and internationally) → vicious cycle

1. **Sex ratio:** improves as the share of mean economic revenue increases. i.e. the higher the mean income level, the higher the gender balance → dependence of poorer households on migrational labour revenue.

1.3. Methodology and Thesis Structure

Upfront, it is important to note that this dissertation is mainly an empirical analyses dissertation. The basic theoretical framework is developed by Ferreira and Gignoux (2011), from now on abbreviated as FG (2011), which is then adapted to the dataset and the Nepalese country specific context.

This study focuses on Nepal and primarily utilises the 2011 Nepal Living Standards Survey Round Three (NLSS III) household (HH) dataset. In order to allow for an optimal model adaptation and adequate result interpretation, a comprehensive overview of Nepal with essential country background information (e.g. history, social structure, geomorphology) is given.

The utilised framework in this study is partially based on the model developed by FG (2011) and the theoretical concept of IOp draws on Roemer (1998). Roemer explains that differences in *outcome* or *advantages* (y) can be explained by firstly, exogenous *circumstances* variables (C) to an individual, such as race or parental background, and secondly, an endogenous *effort* variable (e) for which an individual can be held responsible and accountable. The effort variable can contain elements

such as time spent studying, hours worked, personal ambition, social capital formation, decision taking or aspiration. Perfect equality of opportunity (EOp) exists when outcomes are independently distributed of circumstances and solely depend on efforts.

The main analysis of this dissertation has two dependent outcome variables (y): per capita consumption expenditure and per capita income, and seven circumstance variables: gender, ethnicity, religion, father's and mother's education, and two birth region groupings. Individuals are divided into different *types* or *groups* in which they are homogenous in terms of theirs circumstance characteristics. Individual *i* is characterised by (y_i,C_i^j,e_i).

The empirical analysis mainly draws on two models used by FG (2011). Firstly, a model inspired by Bourguignon et al. (2007), a linear regression and inequality decomposition model based on the following reduced form: y=*f*[C,E(C,v),u]. Secondly, a model inspired by both Bourguignon et al. (2007) and Checci and Peragine (2010), a parametric and non-parametric approach for both absolute (IOL) and relative (IOR) inequality of opportunity indices. IOR measures the level of IOL in relation to total inequality in outcome. These indices exclude the endogenous effort variable and satisfy the standard axiomatic properties in the literature on the measurement of relative inequality. Furthermore, they satisfy the path independence property.

IOR, IOL, various income disparity measures (e.g. Gini coefficient, Mean Log Deviation), and OLS regression estimates are compared. There are four levels of empirical econometric analyses: national, Regional Development, urban and rural area, and economic outcome groupings by population quarter level. The results are checked for robustness and reliability at all stages of the analyses. The IOp indices are additionally checked using the equivalence scale measure.

	Parametric model	**Non-parametric Model**
IOL	$\theta_a^{PS} = E_0(y) - E_0(\tilde{v})$	$\theta_a = E_0\left(\left\{\mu_i^k\right\}\right) = \frac{1}{N}\sum_{i=1}^{N}\log\frac{\mu}{\mu_i^k}$
IOR	$\theta_r^{PS} = 1 - \frac{E_0(\tilde{v})}{E_0(y)}$	$\theta_r = \frac{E_0\left(\left\{\mu_i^k\right\}\right)}{E_0(y)}$

μ is the mean group advantage for *k* type and *l* type, E_0 indicates the inequality index in advantage *y,* and the smoothened and standardised advantage *v*.

These pragmatic and relatively simple scalar measures of IOp also allow for analysing the most opportunity-deprived types and their specific circumstantial profiles for economic outcome. The cross-comparison of these results is of direct relevance for formulating customised and localised opportunity-equalising policies.

This dissertation takes the following **structure**. In order to wholly analyse Nepal for the existence of IOp, this dissertation consists of two main analytical chapters, which are complementary. Chapter 2 investigates the level of inequality in outcome (income and consumption expenditure), which serves as the basis for the subsequent analysis of inequality of opportunity in Nepal (Chapter 3). Chapter 4 compares and highlights the literature contributions of the previous two chapters. Chapter 5 concludes this dissertation.

In more detail, Chapter 2 first discusses the literature on the measure of welfare and the suitability of income and consumption as proxies for measuring economic welfare. Then, a thorough

look at the most conventional literature on the measure of inequality in outcome is given, and the most suitable inequality in outcome measure, the Mean Logarithmic Deviation, is discussed as the most suitable reference index for the subsequent chapter on inequality of opportunity. The used Nepal Living Standard Survey dataset is described in order to facilitate the understanding of the two sections that extensively elaborate on the aggregation (methodology, description, robustness checks, limitations, and extensions) of both consumption and income. Next, the aggregates, their entailed economic outcome distributions and disparities, and the consequently computed inequality indices are analysed at the descriptive level. Finally, some policy recommendations for combatting inequality in outcome are briefly discussed. The chapter concludes and highlights the importance for a more complete inequality measure than *just* inequality in outcome.

Chapter 3, first, gives an overview of the IOp literature, terminology, basic frameworks, underlying principles, challenges, and trade-offs in computing the indicators and the interest for the public and policy makers in the study of inequality of opportunities. Next, the methodology is explained in both an intuitive and a theoretical manner. The country-specific background is given. The dataset is described in more detail, and the applied dataset constraints, the author's attempts in overcoming some of the dataset shortcomings, the choice of the dependent and independent, and the econometric model adaptations to the Nepalese dataset are discussed and justified. Four separate econometric sections follow for the study of IOp: national level, Development Region level, urban rural area level, and population grouping by economic welfare quarter level analyses. Each of these contains a descriptive and econometric analysis, computes the IOp indices and runs the reduced-form ordinary least square regressions, specifies the circumstantial profiles of the most disadvantaged types, and discusses the numerous robustness checks. Then the main results from all four levels of analyses are compared. Some policy implications, study limitations, and further extension ideas are given. The last chapter concludes.

On a side note, Chapter 2 and 3 are complementary for the wholistic understanding of IOp in Nepal. Chapter 3, however, is the main focus of this dissertation, thus its disproportionate length compared to the other chapters. While Chapter 3 could be divided into two chapters (national level analysis and disaggregate level analysis), due to a large overlap in terms of methodology, results, policy conclusions and so on, it was decided to keep the two chapters as one.[1]

Chapter 4 compares and highlights the literature contributions of the previous two chapters on a methodological and empirical level. Chapter 5 concludes this dissertation.

[1] The aim is to publish Chapter 3 as two papers. First, the national level IOp analysis with an emphasis on comparison to the existing literature, most of which also only focuses on national level IOp computation. Second, the disaggregated national level IOp analyses adding to the IOp literature in terms of depth and thorough understanding of IOp with its regional and population differences across one country. This then allows for more customised and effective IOp policy framing. The second paper would thus be a more Nepal contextual specific analysis with important policy implications.

Chapter 2 Inequality in Outcome for Income and Consumption Expenditure: Analyses of Nepal

This chapter first gives a brief literature review on the measure of economic welfare and highlights some of the advantages and disadvantages for the utilisation of income and consumption expenditure as welfare measures. Second, this chapter provides a brief overview of the inequality in outcome literature, focusing mainly on disparity measures and index computations. Third, the chapter then focuses on the methodological procedures to compute both the consumption expenditure and income aggregates and highlights some of the descriptive statistics. Fourth, a descriptive comparison is given between the consumption and income aggregates, between the various distribution and disparity measures, and between the numerous inequality in outcome indices. Finally, the chapter concludes by giving some policy recommendations, highlighting some contributions to the literature, and opening on the need for a more complete inequality measure to complement the standard inequality in outcome measures.

2.1. Literature Review: Economic Welfare, Income and Consumption Expenditure Aggregation

The **literature on welfare**, and more precisely on economic welfare, is abundant at both the normative and positive levels. The study of human welfare is a central pillar of economic research and, thus, has been the source of extensive studies on what welfare actually is and how it can be measured. Dowding (2009) elaborates on this in an entire chapter in the *Oxford Handbook of Philosophy of Economics*, with the basic theoretical line behind human welfare being the goal to maximise the satisfaction of human preferences. In other words, this is the objective to optimise the combination of choices available to individuals (or a group or society) in order to achieve the highest level of welfare possible. The aim of this dissertation is not to summarise the literature on social choice theorems, welfarism, or utilitarianism. Instead, it builds upon the generally agreed and most frequently used monetary economic welfare proxies, namely income and consumption expenditure (hereafter, consumption and consumption expenditure are used as interchangeable terms) (Jenkins and Van Kerm, 2009). These two monetary proxies for living standards are considered to be the easiest monetary measures to calculate and compare across datasets, according to Deaton and Zaidi (2002). Yet they do not capture other, less quantifiable components of living standards, such as freedom and human rights, happiness and fulfilment, or health status and life expectancy. It is, therefore, important to accurately calculate income and consumption expenditure so that it adequately, to the largest extent possible, reflects the economic welfare of a household (HH) or an individual.

Despite the abundance of literature on how to compute income and consumption expenditure, there is little consensus on the "ideal" way of doing so. Some attempts to provide general guidelines on computing these, however, do exist (e.g. Deaton and Zaidi, 2002; Carletto et al, 2007, Quinones et al., 2009). Utmost care must be taken to accurately aggregate the components of both income and consumption in order to reliably analyse the livelihood and socio-economic conditions and choices of HHs and individuals. Section 2.4.1 and summary Table 5.2 briefly summarise the literature on the different methodological guidelines to compute the consumption aggregate. Section 2.5.1 does the same for income. Both sections furthermore suggest an "optimised" aggregation methodology proposed by the author that is adapted to the Nepal Living Standard Survey, third round (NLSS III) dataset.

Neither consumption nor income are perfect or complete indicators for measuring an individual's or a HH's welfare, but they are often the best available ones and the easiest to measure monetarily, and they capture the control over HH goods. Furthermore, usually, Living Standard

Measurement Survey (LSMS) datasets provide ample information to adjust income and consumption to differences in local costs of living. In brief, income tends to measure financial flows, while consumption focuses on the actual utilisation and absorption of goods and services. Both aggregates have their advantages and disadvantages.

The literature suggests that **consumption is a better welfare proxy than income** (e.g. Deaton and Zaidi, 2002; CBS, 2011; Ferreira and Gignoux, 2011; Ray, 1996; Banerjee et al., 2006). First, Deaton and Zaidi (2002) argue that consumption measures a person's actual well-being in terms of meeting his/her current basic needs, while income only serves as an element that allows such consumption.

Second, consumption measures welfare at the current moment in time. It is a cross-sectional portrait of welfare. Income, however, bears both static and dynamic temporal elements (current and future welfare). It is, therefore, also less suitable for the subsequent inequality of opportunity analysis in Chapter 3.

Third, in a developing country context, it is preferable to use consumption over income, because consumption data tends to be more reliable than income data. In brief, this is mainly because income flows tend to be more irregular in developing countries, are often informal and unregistered, and often individuals are rewarded for their services to a large extent with in-kind payments rather than in monetary terms. It is, thus, often more problematic to record income correctly, while consumption represents a monetised accumulation of consumed goods and services, which are comparatively well recorded. In developed countries, income tends to take the regular form of salaries and wages and is, thus, easier to aggregate.

Fourth, consumption is less prone to fluctuate in the short term, even if income fluctuates. So, consumption can be said to be less variable than income. Deaton and Zaidi (2002) argue that consumption reflects a HH's or an individual's welfare better in the long term since consumption smoothens out income fluctuations.

Last, FG (2011) argue that consumption is a relatively reliable reflection of individuals' choices for goods and services and is, therefore, a reasonably adequate representation for individuals' well-being.

Yet, the **consumption measure also bears some shortcomings** that must be taken into consideration, especially when interpreting the econometric results. Deaton and Zaidi (2002), for instance, argue that intra-household resource allocation, i.e. who consumes what, can be difficult to attribute at the individual level. Income components, on the contrary, such as pension or wages, are more easily identifiable at the individual level.

Furthermore, the consumption computation methodology is more complex than the income one. It is thus more prone to computational error and cross-data variation (e.g. country-specific items, such as buffalo meat, chillies, and bitter gourds, or measurement units). This makes the comparison between different country aggregates difficult.

Furthermore, Jenkins and Van Kerm (2009) generalise that the conventional welfarist approach to measuring economic welfare prefers to use consumption, rather than income, as the economic welfare variable for the subsequent analysis of distributional inequality. They argue that this is because consumption is directly reflected in an individual's utility function. However, they continue that this conventional welfarist approach is heavily debated. This is because inequality is often "considered to be about differences in access to or control over economic resources rather than the actual exercise of that power". So from this point of view, income is preferred over consumption.

Given the briefly elaborated trade-offs in utilising either income or consumption as an economic welfare measure, this dissertation uses both aggregates as separate independent variables in the subsequent inequality of outcome and inequality of opportunity analyses.

As the computation methodology sections will discuss, there is no "right way" of aggregating consumption and income. Much of the methodology followed depends on the structure and completeness of the dataset. Key is that the aggregation process must not be done mechanically. For

instance, the extent of inclusion of public goods into either of the aggregates is complicated and debated. In general, it is argued that public goods provision is difficult to measure, and it should thus be excluded. However, under some circumstances it should be included (e.g. when the purpose is to evaluate public good provision). This is discussed further, for instance, in the sections covering the inclusion of public education into the consumption aggregate.

The different guidelines and computation methodologies (e.g. how certain components are weighted, included, or excluded) therefore give different results for the corresponding overall aggregates, and the different aggregates then have different impacts on the econometric analyses. Accordingly, this chapter argues that one must a) pay full attention to carefully clean the data and compute the welfare aggregates and b) fully understand how a variable has been aggregated in order to be able to subsequently interpret the analyses' results.

This chapter elaborates on the computation methodologies of both income and consumption. Both welfare measures are separately used in the inequality in outcome and of opportunity analyses in order to check for robustness (i.e. to complement each other due to their welfare representation shortcomings). Various guidelines are followed to compute the aggregates, and an "optimised" methodology for the NLSS III dataset is developed. The adjusted income and consumption measures then serve for the following inequality analyses.

Given the low development status of Nepal (amongst the 30 poorest countries in the world, according to the World Bank Development Indicators, 2017) and the abovementioned reasons, consumption expenditure is preferred over income as the economic welfare proxy in this analysis. Generally, the consumption aggregates are lower and fluctuate less to those of income. The consequent inequality estimates are also lower.

2.2. Literature Review: Inequality in Outcome. Income and Consumption Expenditure Distribution, Disparity, and Inequality Measures

This section gives a brief theoretical overview of inequality and, more precisely, inequality in outcome. The economic outcome variables utilised are consumption expenditure and income. The main references used in this section are Bourguignon et al. (2007), Bourguignon and Morrisson (2002), Fields (1980), Glyn (2011), Kuznets (1955) and Ray (1998).

The **term inequality** is universally recognised, and the socio-economic issues related to inequality are a source of continuous public debate. Yet, the term inequality is often ill-defined or not defined at all. Some terms, such as distributional inequality or disparity, are frequently used as synonymous to inequality in mass media or by the general public for example, but these terms are quite distinct to the term inequality. Some clarification and definitions follow using the Cambridge and Oxford dictionary definitions. *Distribution* usually refers to the characteristics of the distribution of income, for example, across a population. For instance, one may ask how evenly income is distributed across a population. The focus lies on stating the distributional elements of income.

Disparity, however, usually refers to the comparison of at least two groups and on how they differ within a distribution, i.e. the imbalance or discrepancies in the income levels between them. The emphasis is more on the unevenness or disproportion of the characteristics analysed.

Inequality usually refers to at least two expressions being unequal or different.

Then the questions become: what are the expressions of inequality that are being used and how are they analysed? Does the referred term of inequality have a static or a dynamic element? Are the inequality definitions and evaluations subject to different underlying philosophical and normative thoughts? In other words, for instance, is it income or consumption inequality, access to health care or education inequality, or intergenerational social mobility inequality?

The **concept** of inequality is difficult to define exactly, especially when employing it in a socio-economic context. In general, one says that inequality exists when an individual A has a greater set of choices to choose from than individual B. However, the multi-dimensional aspect of inequality signifies that countless other elements need to be taken into consideration when trying to quantify inequality levels (e.g. do individuals have the same liberties, access to infrastructure and opportunities, or the same biological background).

The most simple and universal measures of inequality are income distribution measures. But even then, numerous debates arise, such as which income distribution measures are the most appropriate or whether one should use the current level income or the lifetime income, etc. Nowadays, most statistical inequality analyses utilise data coming from household surveys, allowing for a HH and an individual level income analysis. When income or consumption are used as the main expressions to analyse inequality, then these are called inequality in outcome measures. This chapter primarily focuses on the analyses of the Nepalese NLSS III dataset for inequality in outcome, i.e. income and consumption.

Inequality is an ever-prominent **subject in the policy environment**. Especially in the developing country context, the trade-offs, relations, and causality of the poverty-inequality-economic growth triangle (Bourguignon et al, 2007) has been subject to heavy ideological, empirical and operational debates. It is generally agreed upon that the study of inequality is a necessity for a political debate in order to elaborate short-, mid- and long-term development strategies. There are some advantages of having inequality. For instance, higher income inequality in general leads to higher savings and thus drives the possibility for capital investment, leading to an upward economic growth spiral. This goes under the assumption that rich people have a higher propensity to save. However, there are also some disadvantages in having high income inequality levels. Often, high inequality is used as an explanation and source for social and political instability, and consequently lower economic growth, when societies become too polarised.

According to Bourguignon et al. (2007), there are two main **reasons** to be interested in the study of inequality. First, **ethical** reasons. What are the given ethical norms in the society that faces income inequality? For instance, should one tolerate the continuation of patrimony and wealth at the expense of treating a segment of the population unfairly from birth?

Second, **functional** reasons. Inequality can have an impact on economic growth, which can have an impact on poverty rates.

This dissertation focuses on the latter of the two reasons, not going into the theoretical and normative debates of inequality in the literature of distributive justice (see Roemer (2009) for a comprehensive summary).This is because this dissertation is mainly an empirical study of inequality in outcome and inequality of opportunity. Opening Pandora's Box of values, ideologies, norms and corresponding theories of distributive justice is out of the scope of this dissertation. Section 3.1.4 briefly touches on the normative versus positive trade-offs.

Glyn (2011) elaborates on two main **distributional differences**. First, he describes the importance for the analyses of functional distribution inequalities. These contain remuneration differences of factors of production, which are independent to the within-household distribution of factors. Second, he elaborates on individual distribution inequalities. This is the individual-level revenue distribution of what households receive depending on their production factor endowment.

There are **various ways to describe and measure income and consumption inequality**. The simplest form is at the descriptive level, e.g. by looking at the distribution and disparity of income. A more sophisticated manner is the numerous inequality indices or measures that take into consideration

the distribution and disparity of income, for instance, to compute one index. It is important to emphasise that different measures place emphasis on different elements and parts of the distribution. An attribute most inequality indices have in common is that the higher the index, the higher the level of inequality. Some range from 0 to 1. Others have no maximum value. The most commonly recognised and utilised indices accept four basic axioms. If all four axioms are satisfied, one can accept the Lorenz criteria, meaning that the Lorenz curves do not cross each other.

The debates on the various **axioms** are important, and for more details Fields (1980) and Glyn (2011) serve as a valuable reference. Here, only the basics are elaborated. It is important to state, however, that the more general the assumptions are, the more easily they are accepted, yet the indices provide different results. The more narrowly defined the axioms are, the more precise and exact the indices are, however, the more debated their utility and validity. The following are the four main axioms in the literature.

A1: Anonymity
A2: Population homogeneity
A3: Principle of relative revenue or revenue homogeneity
A4: Regressive transfer principle

Per capita income, for instance, can be ranked in increasing order. The same can be done for per capita consumption expenditure. Both can also serve as poverty indictors, but by themselves they do not reveal much about the level of inequality. For that, a more complex analysis of income level distribution must be done.

Let $y = (y_1, y_2, \ldots, y_n)$ be the vector that represents the income of a population for which we want to analyse the inequality in income distribution. The inequality index, I(...), is a function that gives each vector y a value of I(y), where $I(y) > I(y')$ only if y is a more unequal distribution than y'. Therefore, the higher the I(y) value the higher the level of inequality.

A1: Anonymity. A permutation between A and B does not affect the index value: $I(y_A, y_B) = I(y_B, y_A)$. For instance, for two populations, where each population has four individuals α, β, δ, and γ:

$$\mathbf{Y_1 = (1,\ 2,\ 3,\ 4)} \quad \mathbf{and} \quad \mathbf{Y_2 = (1,\ 2,\ 3,\ 4)}$$
$$\alpha\ \ \beta\ \ \delta\ \ \gamma \qquad\qquad \alpha\ \ \delta\ \ \gamma\ \ \beta$$

In population Y_2, the incomes are the same as for Y_1, but the incomes go to different individuals. So, if the anonymity axiom is satisfied, the inequality level between the two populations will be identical. This axiom is largely accepted (Fields, 1980).

A2: Population homogeneity. Doubling two populations with the same characteristics will give the same inequality level, $I(y_1, y_2, \ldots, y_n) = I(y_1, y_1, y_2, y_2, \ldots, y_n, y_n)$. For example, take copying the same population and doubling it. Just by doubling it, the level of inequality does not change: $Y_1 = (1, 2, 3, 4, \ldots, 7)$ and $Y_2 = (1, 1, 2, 2, 3, 3, 4, 4, \ldots, 7, 7)$. Therefore, there is a need for a comparison criteria that allows to compare populations of different sizes. The axiom of population homogeneity allows for the comparison of inequality levels between different sized populations. In other words, the number of individuals within a sample does not matter. This principle is also largely accepted (Fields, 1980).

A1 and A2 emphasise the importance of elements or shares of the population and their respective income shares.

A3: Principle of relative revenue or revenue homogeneity. An inequality index is homogenous in income to the degree 0 when $I(y_1, y_2, \ldots, y_n) = I(\lambda y_1, \lambda y_2, \ldots, \lambda y_n)$, $\lambda > 0$. The principle

assumes that the marginal income utility is constant, meaning that it is the relative and not the absolute differences in revenue that matter. For this, one has to accept that the marginal income utility is constant, which raises controversy for the homogeneity in relative income axiom.

Table 2.1: Example Table to Explain the Principle of Relative Revenue (Axiom 3)

Average per capita income	Country A	Country B	Difference in income between country A and B
y_{t1}	20,000	2,000	18,000
y_{t2}	40,000	4,000	36,000

For example, country A has an average per capita income of 20,000 units at t_1 and country B of 2,000. Country A's level is about 10 times that of country B. If over the same time period both countries double their respective per capita income levels, a relative income inequality measure, like the Gini coefficient, will not change. Both countries doubled their average per capita income levels. However, the absolute difference in average income revenue between the countries increased from 18,000 to over 36,000 units. One can compare the income distribution between two populations that have different average incomes.

A4: Regressive transfer or Pigou-Dalton principle. This is one of the most important criteria (Fields, 1980). It is based on a regressive transfer principle where distribution A = distribution B, and which flows from the poorest to the richest. This will increase inequality. The income vector of (y_1, …, y_i, …, y_j, …, y_n) is ranked from the lowest to the highest income earner. The transfer axiom is satisfied if for all transfer δ for all i, j so that $y_i < y_j$ the inequality index $I(y_1, \ldots, y_i-\delta, \ldots, y_j+\delta, \ldots, y_n) > I(y_1, \ldots, y_i, \ldots, y_j, \ldots, y_n)$.

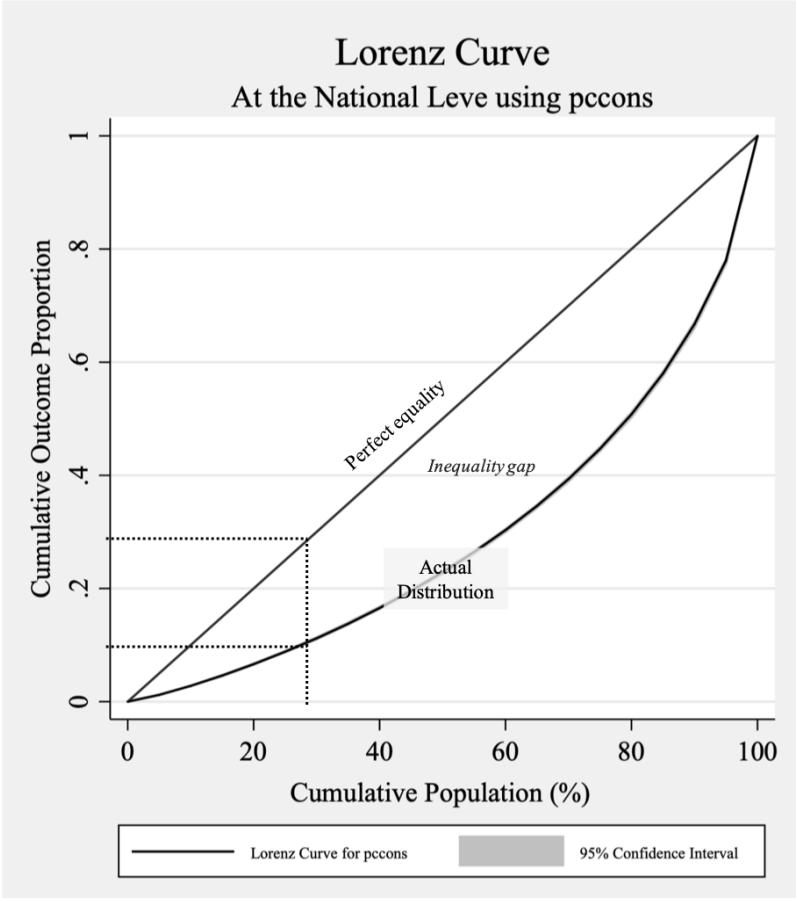
If all four axioms are satisfied, then the **Lorenz criteria** can be accepted. At the most intuitive level, the Lorenz curve is a graphical representation. Figure 2.1 illustrates three different empirically calculated Lorenz curves, including one computed by the author utilising the NLSS III dataset and the adjusted per capita consumption aggregate. The author's computed curve represents cumulative shares of consumption held by the cumulative share of the population. The shares can be divided into deciles, quintiles, or any other way. To the bottom left, the Lorenz curve begins at 0, where 0% of the population holds 0% of consumption. A perfect straight line with a unitary angle (45° angled line) represents perfect equality, where each additional share of the population holds an exactly equal share of consumption. If the curve's angle is initially inferior to one, it means that the share of the population holds a disproportionately smaller share of consumption. For example, the bottom 10% of the population, hold less than 10% of the national consumption level. With inequality, the marginal consumption of each following share segment of the population is inevitably higher than the previous population segment. That is, generally, the angle increases the further one moves to the right on the curve. This means that a relatively small percentage of the population earns disproportionately more than their population share in consumption. Furthermore, the value of the Gini coefficient equals double the area marked as "inequality gap".

The Lorenz curve is necessarily convex and below the 45° line because the population is always ranked from those with the lowest to those with the highest consumption levels. If the Lorenz curve is an exact inverse L shape, then there is perfect inequality. The higher the level of inequality, i.e. the higher the consumed shared of consumption owned by a lower share of the top population, the further away the Lorenz curve is from the original perfect equality 45° line. Should two Lorenz curves completely overlap, then they have the same degree of inequality. However, should two curves cross each other, one cannot conclude under the Lorenz criteria which population is more unequal in terms of the consumption distribution.

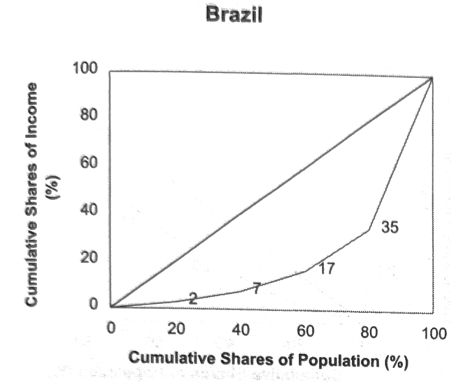
Formally, the Lorenz criteria is met when all four initial axioms are satisfied. This means that for each pair of income distributions (y1, y2, …, yn) and (z1, z2, …, zn) : I(y1, y2, …, yn) ≥ I(z1, z2, …, zn) when the Lorenz curve for (y1, y2, …, yn) is either on the left or on the right of the Lorenz curve of (z1, z2, …, zn).

The above axioms and the Lorenz curve criteria bear information on how the economic outcome (i.e. income or consumption) is distributed with and between populations. However, they do not highlight the absolute difference in distribution. If the Lorenz curves cross each other, one can only highlight the weighted difference given to more or less inequality for the poorest compared to the richest and vice versa. Furthermore, the Lorenz curve does not allow for a complete classification of economic outcome distributions. Therefore, more complete inequality measures are often put forward that summarise inequality levels in one estimate.

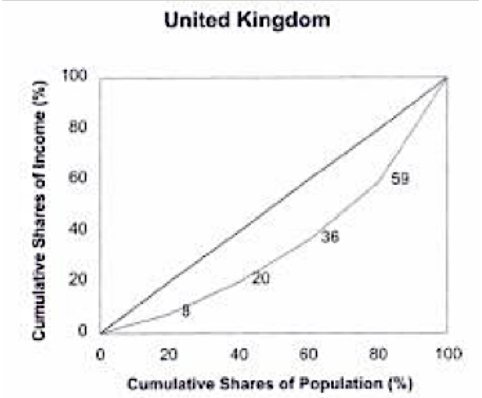
Figure 2.1: Empirical Lorenz Curve Examples



Source: Author's computation



Source: Ray (1998)



Source: Fields (1980)

At the **descriptive level**, there are countless other ways, apart from the graphical Lorenz curve illustration, to look at inequality in outcome. Table 2.2 is a summary table of some of the simplest inequality in outcome measures (e.g. the Rich-Poor Interval, the Absolute Average Variance, the Coefficient of Variation, and percentile ratios). The table states the basic underlying computation equation and gives a brief explanation as well as some advantages and disadvantages of the measures. This dissertation also provides information on distributional inequality by population quintiles (e.g. Table 2.16 and Table 2.17), on distributional differences across the country by geographical criteria (e.g. Table 2.18 through to Table 2.23), as well as on percentile ratios (Table 2.15 shows some selected percentile ratios for the various computed income and consumption aggregates. **Percentile ratios** are

the relative distance between two points considered, depending in the top, the lower or the middle range of the income distribution. Upfront and as expected, the percentile ratios for income are always higher than for consumption. The only exception being the p10/p50 percentile ratio. This is because it measures the ratio of the bottom 10% of the distribution compared to the median i.e. looking exclusively at the lower parts of the distribution (bottom 50%). For this, the consumption value is higher than for income because the distribution of consumption gradually increases and is more equal (also in the lower parts of the distribution) than for income (see the Lorenz curve comparison in Figure 2.5). So, in terms of consumption, an increase in economic welfare can be noted within the lower 50% of the distribution towards the distribution median. For income, however, there is not much difference between the bottom 10% earners and the earners at the median level. This is because at the income level, a) the society is a lot more unequal, where most income is earned by the top few percent earners, and b) a relatively small proportion of the country has regular incomes, Nepal being an underdeveloped country. This means that when only looking at distributional differences of income of the bottom 50% of the population, they are all more or less "equally" bad off (see almost horizontal slope for the lower 50% cumulative population share of the Lorenz curve in Figure 2.5).

For the remaining reported percentile ratios, the income ratios are higher than those for consumption. This also goes for p90/p50 (i.e. focus on the upper 50% part of the distribution), which compares the top 10% earners (or consumers) to the median. Following the argument from p10/p50, the top 10% hold a much larger share of economic outcome than the rest of the country, so that when they are compared to the median, the inequality level of the top 10% compared to the top 50% is very high. And that is all the more important for income earner than for consumers.

The p90/p10 in the table represents the comparison of the 90th to the 10th percentile of the outcome distribution, i.e. how much the top 10% of the population hold compared to the bottom 10%. With a factor of 14.7 for the reference income aggregate, the level of unequal distribution is the highest amongst all given percentile ratios (as expected). It is also about 3 times that of the reference consumption variable.

P75/p25 compares the distributional inequality level of the middle class, suggesting that within group distribution inequality is less when using consumption over income as the outcome variable. This again suggests that income distribution is more skewed towards the top end of Nepalese society, while consumption seems to follow more of a normal, centralised distribution pattern.

Furthermore, a stark difference across all percentile ratios is noticed depending on which aggregates are used. For instance, for the consumption aggregate and the p90/p10 ratio, distribution inequality levels vary from 5.2 when using per capital consumption excluding education and health expenses (monthly criterion) compared to 8.1 when using Deaton and Zaidi's (2002) weekly criterion aggregation methodology. This suggests again that the aggregation methodology used to compute the economic outcome variables is crucial and directly impacts on its distributional inequality.

Table 2.15Table 2.15 shows some selected percentile ratios for the various computed income and consumption aggregates. **Percentile ratios** are the relative distance between two points considered, depending in the top, the lower or the middle range of the income distribution. Upfront and as expected, the percentile ratios for income are always higher than for consumption due to the greater variance of income compared to consumption.

The distributional outcome differences across a population are often analysed by population quarters, quintiles, or any other population grouping. The **population** is divided into equally large segments of the population, and a comparison is made between their respective shares of overall economic outcome.

Giving the description of inequality in outcome a **geographical dimension** can provide further value for the analysis of inequality in a country. In this dissertation, three main geographical categories are used (for a complete elaboration on this, see Section 3.3.1). 1) Nepal is divided into three ecological

regions: the high mountain region, the central strip of the hill region, and the low-lying flatlands of the Terai region. 2) The data is also divided into urban and rural areas. 3) The country is administered by five publicly divided Development Regions: the far-western, mid-western, western, central, and eastern regions.

Table 2.2 provides a summary of the use of **percentile ratios**. These ratios are the relative distance between two points considered, in the top, the lower, or the middle range of the outcome distribution. For instance, a p50/p10 ratio examines distributional inequalities in the lower half of the income distribution compared to the bottom 10% of the distribution. It is the equivalised median income to the 10th percentile income level. Public debates often refer to the top 1% of a population earning X% of the national overall income in order to highlight the significant imbalance in earnings (e.g. p99/pX).

These various descriptive measures bear valuable information for the analysis of inequality in outcome distribution and disparity. Utilising income and/or consumption as a measure for welfare, they highlight the degree of their respective distribution, disparity, and inequality. Yet, more complete inequality indices are often used to complement the descriptive level analyses. Summary Table 2.3 and Table 2.4 for the more complex inequality measures seek to complete the summary table for the simplest inequality in outcome measures. It is important to highlight that choosing one inequality concept over another depends on the choice and judgement of an ethical and normative underlying basis.[2] Thus, each inequality measure places emphases on different parts of the income distribution.

Following are three summary tables for the most frequently used inequality in outcome measures. They state the main computation equations and variable definitions, a verbal summary, and some advantages and disadvantages for each index. The following notations count for all the following inequality measures in the summary tables.

- Number of income groups m, where j: (j=1,… m)
- Number of individuals per group n_j that earn an income y_j for group j
- The total population is the sum of individuals in each group N= $\sum_{j=1}^{m} n_j$
- The average income μ of the population is $\mu = \frac{1}{n}\sum_{j=1}^{m} n_j\, y_j$

[2] Note that it is not within the scope of this dissertation to question and analyse the underlying norms.

Table 2.2: Comparison Table for the "Simplest" Inequality in Outcome Indices

Income inequality measure	Rich-poor Interval (R)	Average Absolute Variance (M)	Coefficient of Variation (CV)	Percentile Ratios (PR)
Computation equation(s)	$R = \frac{y_m - y_1}{\mu}$	$M = \frac{\sum_j n_j \lvert y_j - \mu \rvert}{n\mu}$	$CV = \frac{\sqrt{\sum_j n_j \left(y_j - \mu\right)^2}}{n\mu}$	$PR = \frac{Q^i}{Q^i}$ where Q^i is the i's percentile of the income distribution (Fessler et al., 2013)
Summary	Only takes into account part of the population and distribution. It is the relative gap between the richest and poorest. Only the gap is observed.	Takes into account the entire population and income distribution. The larger the difference from the average income, the higher the inequality level. I.e. inequality is proportional to the gap with the average income level.	Takes into account the entire population and income distribution. It is the standard deviation divided by the mean.	Percentile ratios give a first overview into income distribution inequality. They measure the relative distance between two points considered, in the top, the lower or the middle range of the income distribution. For instance, a p50/p10 ratio examines distributional inequalities in the lower half of the income distribution. It is the equivalised median income to the 10th percentile income level. Any other possible ratio combination can be computed. For comparisons of the lower-end income distributions, they can be,, for instance,, p50/p40 and p50/p1, and for the top-end distributional comparison, for instance, p99/p50 and p75/p50.
Advantages and disadvantages	Gives a first aperçu of the distribution levels. Only looks at the absolute or relative gap between the richest and poorest.	Inequality is proportional to the distance of the mean. Does not satisfy the Pigou-Dalton axiom, i.e. a regressive transfer between individuals who are both above the average income has no impact on the indicator.	Easy to understand and immune to outliers. Gives increased weight to more extreme values (squared value). Satisfies the Pigou-Dalton axiom and all Lorenz curve criteria. CV increases when there is a regressive transfer. When Lorenz curves cross each other, the Gini and CV can give conflicting information about their evolution.	Easy to calculate. Jenkins (2009) argues that a particular advantage is that for the p90/p10 ratio, it avoids problems of "top-coding" in survey data. Yet, a major disadvantage being that given the nature of percentile ratios, they ignore any other information apart from the distributional information for the selected percentile. According to Voitchovsyk (2005), the measure is not perfect and especially not for between country ranking, since countries rankings may change significantly when different ratios are used. She also argues that mismeasurement and the influence of extreme values can have a significant impact on the values.

Table 2.3: Table of Comparison for the Atkinson and the Generalised Entropy Indices

Income inequality measure	**Atkinson Indices A(ε)**	**Generalised Entropy Indices G(α)**
Computation equation(s)	$A(\varepsilon) = 1 - \left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i}{\mu}\right)^{1-\varepsilon}\right]^{\frac{1}{(1-\varepsilon)}}, \varepsilon \neq 1$ $A(\varepsilon) = 1 - \frac{\prod_{i=1}^{N}(y_i^{(\frac{1}{N})})}{\mu}, \varepsilon = 1$	The general G(α) formula: $G(\alpha) = \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i}{\mu}\right)^{\alpha} - 1\right]$
Summary	Most of this computation is done as suggested by Jenkins (2008). Key literature and debates are based on Atkinson (1970) and Blackorby et al. (1981). ε measures the aversion to inequality, i.e. it is the inequality aversion parameter. A higher magnitude of ε implies a higher degree of inequality aversion (i.e. more sensitive to income differences in the bottom part of the distribution). The higher the differences in income for the lowest segments of the population, the more positive ε is. Places additional assumption on the social evaluation functions (less constringent) than the Generalised Entropy indices that use an entire spectrum of axiomatic properties to characterise their indices. Yet, the Atkinson indices satisfy axioms 1 to 4 and (depending) also the transfer sensitivity axiom.	Most of this computation is done as suggested by Cowell and Kuga (1981), Jenkins (2008), and Jenkins and van Kerm (2009). Key literature and debates are based on Shorrocks (1984). A comparison of the literature on G(α) and A(ε) is given by Cowell and Jenkins (1995). The more positive α is, the more sensitive the generalised entropy index is to differences in income, particularly for the richest segment of the population. The same goes vice versa, i.e. the more negative α is, the more sensitive G(α) is to a variation for the poorest groups of the population. The estimates of G(α) can reach from 0 to 1 or ∞. The higher the estimate, the higher the degree of inequality, with emphasis on different parts of the distribution depending on the value of α used. G(0): represents the Mean Logarithmic Deviation. G(1): represents the Theil Index. G(2): represents half the square of the Coefficient of Variation.
Advantages	Single-digit indices that allow for cross-country comparison.	Single-digit indices that allow for cross-country comparison. They are amongst the most accepted and recognised inequality in outcome indices. They have been the source and key reference for many further studies of inequality. The Generalised Entropy indices are decomposable by sub-groups after running the first main commands; the Atkinson indices are not (Jenkins, 2008).
Critiques	Emphasis is mainly placed on the lowest segment of the distribution. Assumptions for the Atkinson function are sometimes considered too weak. The Atkinson indices are not decomposable once their index values have been computed.	Each of the five main indices of the Generalised Entropy have their advantages and disadvantages. More details are given later, particularly for the Mean Log Deviation index (G(0)).

Table 2.4: Table of Comparison for the Main Inequality of Outcome Indices of the Generalised Entropy

Income inequality measure	**Gini Coefficient (G)**	**Mean Logarithmic Deviation (MLD)**	**Theil Index (T)**
Computation equation(s)	$G = \frac{\sum_j \sum_k n_j \, n_k \left\lvert y_j - y_k \right\rvert}{2n^2 \mu}$	When α takes 0, giving G(0) in the Generalised Entropy Index computation. $G(0) = \frac{1}{N} \sum_{i=1}^{N} \ln\left(\frac{\mu}{y_i}\right)$	N_k = sample size of group k. $T = T_{w(within)} + T_{b(between)}$ $T = T_w + T_k$ $T = \Sigma_k \theta_k \left[\ln\left(\frac{Y_k/Y}{N_k/N}\right) + T_k\right]$
Summary	Results from 0 to 1, where 0 represents perfect equality (i.e. each part of the population earns a representatively proportional part of the national revenue) and 1 perfect inequality (i.e. the top earns 100% of the national revenue). Takes into account the entire population and income distribution. Takes the differences between all income pairs and sums the absolute value of all the differences (as if it is the level of "inequality between two people"). A regressive transfer increases the Gini coefficient. Fulfils the Lorenz curve criteria. - n^2 pairs - divided by 2 to avoid double counting of the pairs - sum of all grouped pairs of income $\left\lvert y_j - y_k \right\rvert$ normalised or weighted by the groups $n_j n_k$ The Gini coefficient is the most sensitive to differences in income for the income groups around the middle of the distribution (mode) according to the suggested computation by Jenkins (2008).	G(0) of the Generalised Entropy Indices represents the Mean Logarithmic Deviation or MLD. MLD is a logarithmic value of the income distribution and is particularly sensitive to distribution inequalities in the lower and middle part of the distribution.	Probably the most used index of the generalised entropy indices. Results from 0 to 1, where 0 represents perfect equality and 1 perfect inequality. The Theil index is the sum of within and between group inequality. Half the square of the coefficient of variation. T_w is the weighted within group inequality average and T_b the average between group average. The Theil index belongs to the group Generalised Entropy Indices, where α takes 1, giving G(1).
Advantages	Allows for national income distribution comparison. It includes all distribution data. The Gini coefficient is most sensitive to income differences about the middle (more precisely, the mode). Differences of revenue pairs and sum of absolute value of these differences. Existence of both a relative (here) and absolute Gini index. Other Gini adaptations exist e.g. Gini of education.	The biggest advantage compared to the other GE indices is its path independence property. Another advantage is its decomposable between income sources and between individuals. An individual's income is compared to the mean. It does not have a maximum level for inequality.	Advantage that it can be decomposed into between- and within-group inequality. Utilisation of group data.
Critiques	Gini coefficient barely changes over time. When Lorenz curves cross each other, the Gini and CV can give conflicting information about their evolution. The Gini coefficient is not decomposable.	Depending on the normative background of the analysis, the MLD can be considered as placing too much emphasis on the bottom parts of the distribution.	Not directly comparable across populations of different sizes and group structures. Complicated to compute and sensitive to transfer in lower parts of the distribution.

The tables above summarised the key inequality in outcome measures in the literature. They are the most basic ways to analyse distribution and disparity of any given group of individuals. Salverda et al. (2009) give an overview of the theory and particularly the empirics of these indices in their *Handbook of Economic Inequality*. Section 2.6 of this dissertation provides the results for a number of the above-discussed distribution, disparity, and inequality in outcome measures using the NLSS III dataset.

The **Mean Logarithmic Deviation** (MLD or G(0)) is one of the main indices of the Generalised Entropy indices, where α takes 0, giving G(0) in the Generalised Entropy Index computation. The MLD bears numerous advantages over the other inequality in outcome measure for the subsequent inequality of opportunity analyses. First, the MLD is a logarithmic value of the income distribution and it is particularly sensitive to distribution inequalities in the lower and middle parts of the distribution (i.e. supposedly the most opportunity deprived people in the population). This is important for the subsequent IOp analyses, since it tries to study the economic welfare and exogenous characteristic of the most opportunity-deprived individuals or groups who thus face the lowest levels of economic welfare.

Second, the main advantage of the MLD over all other GE inequality indices is that it follows the property of path independence when computing an arithmetic mean (e.g. Roemer and Trannoy, 2015; Ramos and Van De Gaer, 2012; Shorrocks and Wan, 2005; Foster and Shneyerov, 2000). Foster and Shneyerov (2000) give an comprehensive theoretical justifications behind it, and discusses the advantages and disadvantages of it for measuring inequality. They refer to the term as *path independent decomposability.* If the inequality decomposition satisfies the path independence property, then the non-parametric direct and indirect (that is between and within type) approaches provide the same and reliable results (Ramos and Van De Gaer, 2012). More on this in the IOp model specifications in Section 3.2.3.

Third, a big advantage of the MLD is that it is decomposable between income and consumption sources and between individuals, groups, and sub-groups. It is the latter that is of particular interest for the IOp study, which divides the population into "types" or groups of individuals in which they are homogenous in terms of circumstances. A full elaboration on this can be found in Section 3.1, which gives a detailed literature review of IOp, and Section 3.2, which describes and explains the IOp methodology applied to the NLSS III dataset in this dissertation.

Fourth, an individual's income (or consumption) is compared to the mean of the population. Nepal is an underdeveloped country where the national mean is incredibly low and the vast majority of the population faces similar economic burden. Yet, the slightest change in economic gains or loss for the most impoverished can signify a relatively important improvement or deterioration in welfare. Thus, using the mean as a reference is favourable.

Fifth, the MLD does not have a maximum level of inequality. The indicator ranges from 0 to infinity. This allows for upward flexibility when geographically or demographically decomposing the indicator.

Last, the key literature (Ferreira and Gignoux, 2011) on which the IOp methodology of this dissertation is largely based also utilises the MLD as the reference inequality in outcome variable. Since this study only focuses on the empirical analysis of IOp for Nepal, it is of interest to increase the comparability of its results to those in the literature, namely those of the Latin American countries analysed by Ferreira and Gignoux (2011).

For all these reasons, the MLD is used as the main reference inequality of outcome measure for the IOp analyses in this dissertation.

Furthermore, Table 5.3 in the appendix summarises the main empirical IOp literature. It also contains information on, for instance, the reference inequality in outcome indices used by each study. As can be seen, a majority of studies use the MLD, but also many studies used, for instance, the Gini index or variance (CV).

This section highlighted the difficulty in defining inequality. It then continued to use income and consumption as the most universally recognised point of reference to a) measure economic well-being and b) use the analysis of its distributional realities in both a simple descriptive and a more advanced statistical manner. Emphasis was placed on the differences between indices due to their respective weighing on different parts of the distribution, depending on the underlying ideologies and norms. Finally, some reasons were given as to why the MLD is used in this dissertation as the main inequality in outcome indicator for the subsequent inequality of opportunity analyses.

2.3. NLSS Dataset: Description and Other Relevant Variables to Aggregate Income and Consumption Expenditure

This dissertation uses the Nepal Living Standards Survey (NLSS) III. The NLSS III is the third nation-wide household survey, and was conducted from February 2010 through to February 2011 (NLSS I: 1995/96), NLSS II: 2004/2005). It is a typical Living Standard Measurement Survey (LSMS), co-financed and technically supported by the World Bank (WB), it also follows the WB's guidelines and methodology for data collection. The corresponding survey reports were published, and the datasets were made publicly available for purchase in 2011.

The NLSS III surveyed 5,988 households and 34,146 individuals across the country. Refer to Central Bureau of Statistics (CBS; 2011, Vol. 1, p. 4-19) for more details on the sampling methodology, which has allowed the data to be nationally representative. The weight variable is included in the dataset. According to the CBS, a post-stratification adjustment was done to allow for seasonal balance.

The raw data is utilised for the computation of income and consumption expenditure for all households and individuals. Numerous different sections from the survey serve as a reference for the computation of these variables. The weight and stratification variables are used to allow for both income and consumption expenditure to be nationally representative.

Section 3.3 provides further details on the dataset and its variables and provides some country-specific background information.

2.4. Consumption Expenditure Aggregation

This section first elaborates on the methodologies used to compute the consumption expenditure aggregate using the NLSS dataset. It then highlights the various robustness checks. Finally, some limitations and further improvement possibilities are stated for the consumption expenditure aggregate.

2.4.1. Methodology: Guidelines for Aggregating Consumption Expenditure Using the NLSS Dataset and Some Descriptive Statistics

There is no consensus on the "ideal" way to compute consumption expenditure. Various guidelines are followed, and according to them different consumption expenditure aggregates are computed. Also, the author develops an "optimised" methodology for the NLSS III dataset. Table 2.5 summarises the different consumption aggregates, with over eight total HH level and eight per capita consumption aggregates. From now, the description and explanations for the consumption aggregate focus on the author's "optimised" computation after having cross-compared the different guidelines and studied the dataset. An overview and comparison of the different methodologies can be found in Table 5.2 in the appendix. It compares the four main strands of guidelines:

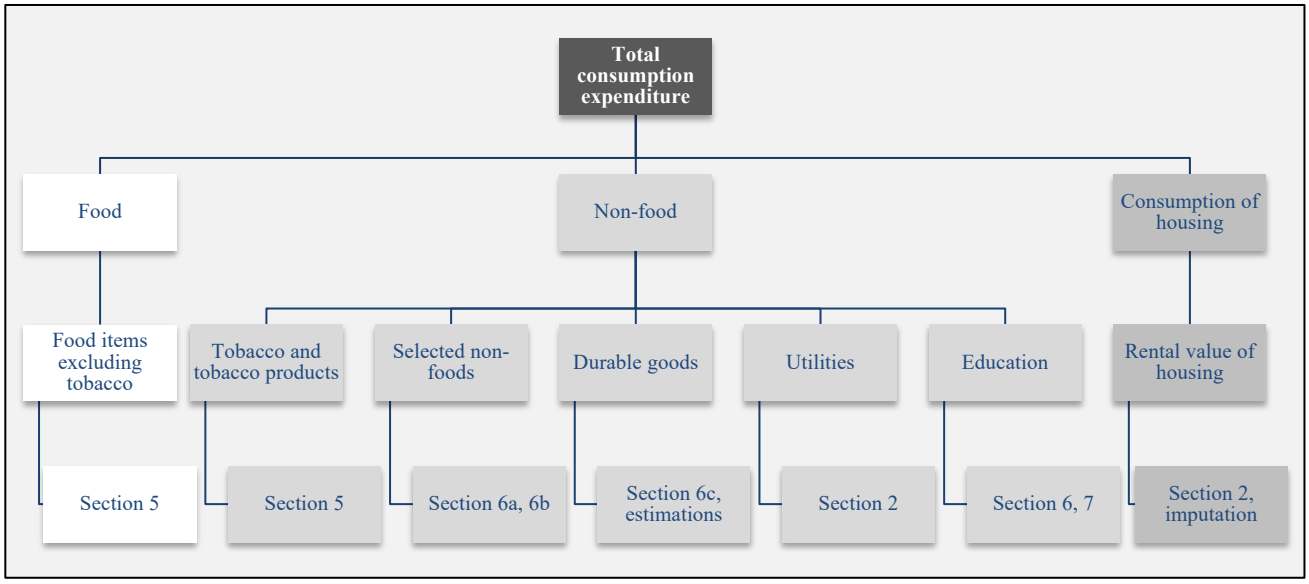
1) Deaton and Zaidi's (2002) guidelines for constructing the consumption aggregate for welfare analysis using the NLSS I dataset. The author replicates this for NLSS III.
2) CBS' (2011) description on how the Nepalese government computed the consumption expenditure aggregate using the NLSS III dataset.
3) Furuta's (2016) elaboration on how to compute consumption expenditure using the NLSS II dataset. The author replicates this for NLSS III.

The author's adjusted method for computing consumption expenditure using the NLSS III dataset. Figure 2.2 illustrates the various components of the consumption expenditure aggregate according to the CBS (2011). There are three main components:

1. Food (72 food items, 13 categories)
2. Non-food (including durable goods)
3. Consumption of housing

These contain various sub-components. The relevant sections of the survey are indicated to facilitate the possible replication of the computation process and to highlight the complexity of combining different elements of the survey to compute the consumption aggregate.

Figure 2.2: Components of Consumption Expenditure Aggregate (Main Components, Sub-Components, Relevant Sections in Survey)



Source: Adjusted from CBS (2011, Vol. 2 , p. 25)

In general, the survey questionnaire asks respondents "How much would you spend in a typical month/past seven days, etc on…" (NLSS III Report 2; CBS, 2011, p. 25-27). Depending on the answers, monthly and then annualised values were computed for each consumption component. This is what the literature refers to as the **"typical month"** or the **"past week" criteria**. In the NLSS III dataset, not always both options are given. Some products are even only asked for on an annual basis. According to Deaton and Zaidi (2002), usually aggregates using the typical week criterion are lower than those using the typical month criterion. This is partially because less frequently consumed and seasonally consumed items are not recorded (e.g. rice bulk buying or school equipment).

The CBS (2011) consumption aggregates are based on the past seven days criterion when available. It is argued that there is an "inherent vagueness in the definition of the "typical month" and that the data suffers from significant recall bias. So, the main advantages of the typical week criterion are that there is less recall bias and there is a specific time horizon. Despite failing to capture seasonally consumed items, the CBS argues that the fieldwork across the country was divided randomly across seasons, which should in turn allow for the data to be representative. It is important to note that the first two NLSS rounds used the typical month criterion only and not the weekly criterion like the NLSS III.

Deaton (1997), on designing and analysing HH surveys, concludes that nowadays, the "typical month" criterion is less commonly used than it used to be, and that there are always trade-offs between using one or the other.

This dissertation therefore computes consumption expenditures based on both, 1) using the typical month, and 2) using the weekly criterion when available. Table 2.5 shows the various consumption expenditure aggregates and their corresponding variable names. The preferred adjusted aggregate is *pccons*.

Table 2.5: Consumption Expenditure Variable Names Following Various Computation Methodologies by Deaton and Zaidi (2002), CBS (2011) Report, Furuta (2016) and the "Optimised" Aggregates by the Author

	Variable name *
The author's computations for NLSS III dataset	
Total household consumption expenditure, typical month criteria, Scholl adjusted aggregation methodology	cons
Per capita consumption expenditure, typical month criteria, Scholl adjusted aggregation methodology	**pccons ****
Total household consumption expenditure, weekly criteria, Scholl aggregation methodology	consWEEK
Per capita consumption expenditure, weekly criteria, Scholl aggregation methodology	pcconsWEEK
Total household consumption excluding education & health	cons2
Per capita consumption excluding education & health	pccons2
Total household consumption excluding education & health using past week criteria	consWEEK2
Per capita annual consumption excluding education & health using past week criteria	pcconsWEEK2
Following Deaton and Zaidi's (2002) guidelines for NLSS III dataset	
Total household consumption expenditure, Deaton and Zaidi (2002) aggregation methodology	consDZ
Per capita consumption expenditure, Deaton and Zaidi (2002) aggregation methodology	pcconsDZ
Total household consumption expenditure, Deaton and Zaidi (2002) aggregation methodology, weekly criteria	consWEEKDZ
Per capita consumption expenditure, Deaton and Zaidi (2002) aggregation methodology, weekly criteria	pcconsWEEKDZ
Following the CBS's (2011) guidelines for NLSS III dataset	
Total household consumption expenditure, NLSS III Report aggregation methodology	consNLSS
Per capita consumption expenditure, NLSS III Report aggregation methodology	pcconsNLSS
Following Furuta's (2016) guidelines for NLSS III dataset	
Total household consumption expenditure, Furuta (2016) aggregation methodology	consFURUTA
Per capita consumption expenditure, Furuta (2016) aggregation methodology	pcconsFURUTA

* Where *totcons* is total household consumption and *pcpcons* is per-capita annual consumption. Both utilise the "typical month" criteria. *totconsWEEK* is the total household consumption utilising the "past week" criteria while *pcpconsWEEK* is the per-capita annual consumption using the "past week" criteria".

** The variable *pccons* follows the "optimised" aggregation methodology as proposed by the author.

Dealing with **outliers** is important in any data cleaning process. The CBS (2011) comments on outliers, but only for the income and not consumption expenditure aggregate (CBS, 2011, Report 2, p 41): "One percent each of both ends of the per capita income distribution -- 120 households in total, 2 percent of all households -- are defined as outliers and excluded from the analysis. All subsequent tables are based on data without the outliers." However, no reasons as to why this threshold is chosen are given. Furuta (2016, p. 165) states that when interviewing Nepalese delegates who conducted and analysed the NLSS II survey data, 0.5% of the top and the lower end of the income and consumption were excluded. This is because their values were deemed too far from the mean and median. In personal exchanges with CBS employees, the author was told that the same approach as with the income aggregate (i.e. the top and bottom 1% rule) was applied to the consumption aggregate. This appears to be a very mechanical approach, whatever the motivation. So, this dissertation follows the advice of various scholars on how to deal with outliers when cleaning data and computing aggregates (e.g. Wooldridge, 2013; Deaton and Zaidi, 2002; Filzmoser and Temple, 2016).

Filzmoser and Temple (2016) and Deaton and Zaidi (2002) state that one should be concerned about outliers because

(a) extreme values of observed variables can distort estimates of regression coefficients;
(b) they may reflect coding errors in the data, e.g. when a decimal point is misplaced or the failure to declare some values as missing; and
(c) they may be a result of model misspecification: variables have been omitted that would account for the outlier, or the outlier may belong to a different population than the one you want to study.

It is thus important to check the data for outliers, miscoding, and the misunderstanding of units, which in turn can affect the unit values. Wooldridge (2013) gives a rule of thumb suggesting to include outliers when using mean values and to exclude outliers when using average values. Wooldridge further suggests that ad-hoc data cleaning should be limited since it decreases the random selection of observation and thus data selection bias.

Furthermore, a more precise look at individual outliers can help to conclude whether the values are even plausible or not. In this case, an empirical approach should be chosen over a theoretical approach. One must consider to what extend an outlier affects the overall desired aggregate and then the overall econometric analysis. In this dissertation, every single component and sub-component and item is checked for outliers, which are then individually analysed. Once outliers are identified and are cross-checked with other HH and individual characteristics (e.g. HH size or proportion to other consumption items), they are either dropped or, where possible, replaced with a plausible corresponding value.

Checking missing prices or unit values is important. The first choice should be the price (unit value) reported by the household. If this is not available, then a proxy should be used. This can be the median (not mean) or the prices paid by "similar" households in the neighbourhood.

Deaton and Zaidi (2002) elaborate on various general issues in computing the consumption aggregate. First, **work-related expenses** should only be included to the extent possible, while purely work-related expenditures should be excluded (e.g. uniform or work commuting costs).

Second, **leisure time** is an important element of consumption and contributes to welfare or the lack thereof. Normally, one would argue that the more leisure time one has, the higher the welfare. Yet, it is difficult to quantify. What is considered leisure and what is each individual's utility function for the consumption of welfare, and to what extent is this quantifiable in a comparative manner? Furthermore, it is difficult to distinguish between leisure, non-market work for the household, and involuntary unemployment. Due to all these problems, leisure time is omitted from the consumption aggregate.

Third, all sorts of production costs **should not be included** (i.e. fertilisers for homegrown tomatoes). This would lead to double counting. Also, all kinds of paid taxes, purchases of assets, repayments of loans, and expenditure on durable goods and housing should be excluded from the aggregate. However, to the extent that local property taxes bear a relation to services rendered, the inclusion of them is recommend. Taxes are deductions on income (flow of money), which do not allow for additional consumption. Since the consumption aggregate tries to measure economic welfare through what is actually consumed by the HH and individuals, whatever means has made it possible means that it is post-tax (i.e. however much disposable income allows for consumption). So, overall the consumption estimate can be interpreted as a post-tax aggregate. Furthermore, savings should not be included since they do not generate welfare through consumption at the time of analysis but are seen (similarly to taxes) as a deduction on income, but with a dynamic dimension. *Regrettable necessities*, such as work clothes, should be excluded from the consumption aggregate but are often too difficult to isolate. Therefore, if there is doubt, they should be included.

Fourth, public goods should not be included since their valuation is difficult. Also, the vast variance in terms of quality and geographical provision make it challenging to include them. Furthermore, some HH may choose to consume private services over public ones (e.g. education). However, housing and health care, if publicly subsidised or provided, can potentially be included. There are more details below in the relevant consumption component elaboration.

Fifth, marriages, dowries, funerals, and similar lump expenditures should be excluded, mainly due to their lumpy characteristics and short time horizons.

Last, theory suggests that gifts, charitable contributions, and remittances provide as much welfare as resources used for something else. Yet, the problem of double counting arises, and they should thus be excluded.

The first consumption component is **food**. All food items are included, except tobacco, which is included in the non-food consumption category. The NLSS III dataset has information on over 73 food items. Food consumption from all sources are included, such as food purchased at markets, consumed outside the house and at restaurants, and received in-kind as gifts or payments or through own cultivation (Deaton and Zaidi, 2002, p.27). The **Engel Law** states that the share of food expenditure decreases as income levels rise (see Figure 5.1 for an intuitive graphical representation). Given that Nepal is one of the 30 poorest and most underdeveloped countries in the world, food consumption takes on a large proportion (41%) of the overall consumption expenditure aggregate for a large share of the population (see Table 2.21: Component Shares of Per Capita Consumption (in %) by Geographical Region)). The World Bank (2016, p. 4) states "Nepal is at a very early stage of development and that the engine of growth has not started cranking up to the extent where this growth acceleration would begin to generate a widening consumption distribution". Thus, particular care must be taken to accurately calculate the food component.

Table 2.23 shows the component shares of food for the per capita consumption aggregate. It indicates that at the national level, over 64% of food items are purchased, over 31% are homegrown or produced, and over 4% are received in-kind. Deaton and Zaidi (2002) highlight the importance to include all food consumption. Numerical values for the **home produced and received in-kind items** are computed. If details for these are not included in the consumption aggregate, they argue that any consequent analysis is likely to overestimate poverty and inequality. Especially given the Nepalese context, where food home production and food received in-kind accumulate to more than 35% of overall food consumption, particular care is taken in calculating accurate consumption prices. In general, Deaton and Zaidi (2002) recommend using a survey respondent's valuation of the goods if they had sold or purchased them at a local market. This should take into consideration the likely difference in quality between home produced and market purchased goods. Respondents' valuation is given for both the typical month and the past week criterion. Both are separately computed and added to the corresponding consumption aggregate.

In-kind received foods include any food products that are received as payments for rendered services, presents, gifts, remittances, etc. In-kind valuation is only given for the typical month criterion, and this also relies on the respondent's valuation of the product had it been purchased (or sold) at a local market. The value is annualised and included into the aggregate.

All food items and tobacco were checked for outliers. One outlier was identified in the "in-kind received" category. The HH's in-kind value was over 300% of the next-highest HH's in-kind received value, while other food consumption elements were within the normal range. The outlier was removed (i.e. one HH of five individuals).

For food consumption, two versions were computed according to the monthly and the weekly criterion. For the weekly criterion, some food consumption observations were at 0, which were judged possible. There is a divergence in the weekly and monthly estimates; however, for the reasons elaborated above, this is considered normal.

The second main component for the consumption aggregate is **non-food expenditure**. The sub-categories include tobacco and tobacco products, selected non-food items, durable goods, utilities, education, and health expenditures.

Selected non-food items are divided into frequent and infrequent expenditures. Both past month and past 12 months expenditure values are given for frequent non-food items. The monthly value is annualised and compared to the annual estimate. They are cross-checked and the most appropriate value is chosen. Infrequently selected non-food expenditures are estimated by the respondent for the past 12 months. The same survey section provides information on some frequently consumed health and education items, as well as fuels. These are dropped here and included in the corresponding categories. Fuels figure in the utilities component. **Tobacco and tobacco products** are included in the selected non-food items component, and information is provided both on the typical month and past week basis.

The inclusion of **durable goods** is based on the rental (usage) equivalent of the product. For instance, instead of including the lump sum purchase price of a new radio or heater, the annual rental equivalent is computed. If the product was purchased and used for the first time over the past 12 months, then the annual rental equivalent without a depreciation rate is calculated. If the product was purchased more than 12 months ago, then various depreciation rates are applied. The depreciation rates are calculated based on how long ago and for what price the product was purchased and for what price it could be sold at the time of the survey interview. The valuation estimates are given by the respondents. Table 2.6 compares the median depreciation rates for durable goods calculated by the author and by the CBS. As can be seen, the estimates are relatively similar, but slightly different. The reason for this is that this dissertation utilises the approach utilised by Deaton and Zaidi (2002, p. 35), where the depreciation rate for each good is calculated in current value terms, and any outliers are individually analysed and eliminated. One outlier is identified and dropped for durable goods older than 1 year, and one outlier for the past 12 months' consumption is identified (motor car) and dropped. The CBS' approach also calculated the depreciation value for each good and then applied the current value. However, at this level, there was no mention of outlier checking.

Table 2.6: Comparison of Median Depreciation Rates for Durable Goods, Computed by the Author and by the CBS (2011)

Durable good item	Median depreciation rate (Scholl)*	Median depreciation rate CBS (2011)**
Bicycle	0.267	0.265
Camera (still/movie)	0.229	0.227
Computer/Printer	0.239	0.242
Fans	0.254	0.253
Heaters	0.294	0.311
Motor car, etc.	0.141	0.146
Motorcycle/scooter	0.197	0.202
Pressure lamps/petromax	0.231	0.228
Radio/cassette/CD player	0.276	0.278
Refrigerator or freezer	0.196	0.187
Sewing machine	0.142	0.124
Telephone sets / cordless phone	0.338	0.351
Television/VCR/VCD Player	0.214	0.208
Washing machine	0.207	0.197

*Depreciation rate calculation results by the author.
** Depreciation rate calculations results by the CBS (2011).

Utility expenditures should be included in the consumption aggregate. They can, however, be problematic since access to utilities is often unequally distributed across a country, some utilities are subsidised in some areas and for some households only, and there can be variations in standards and quality. It is, therefore, difficult to make an accurate regional and international welfare comparison from the pricing of the reported expenditures. Utilities are usually categorised into regularly and non-regularly consumed utilities. Table 2.7 shows the categorisation for frequent non-food expenditure items, including fuel utilities, as advised by the CBS.

Regular utilities include, amongst others, kerosene, matches, and candles. Information is provided with the typical month criterion. No outliers were found, and their values are annualised and included into the consumption aggregate.

Irregular utilities include, amongst others, coal, charcoal, and cylinder gas. Information from the estimated annual values is used, and no extreme outliers were found. These are also included in the consumption aggregate.

Other utilities, such as electricity, garbage, and telephone costs are found in a different section of the survey, and again no outliers were found.

The inclusion of **water** into the consumption aggregate is heavily debated in the literature. Two main arguments being the disparity in quality and geographical access to water, as well as the blur between utilisation for private consumption or for production. First, given the Nepalese context, mainly urban centres have access to clean drinking water. However, urban citizens have to pay for this. While public water purification is scarce in remote, rural parts of Nepal, individuals are either subsidised to access clean water or they do not pay at all. Second, from the survey data, it is difficult to isolate whether consumed water was used for private consumption, such as drinking water, or whether it was used for production, such as crop irrigation. If the latter is the case, it would lead to double counting. In other words, the irrigation price for the water, as well as the home-produced value of the irrigated crop, would be included. These are some of the reasons why water consumption is excluded from the consumption aggregate. The CBS also excludes water.

Table 2.8 briefly summarises the main utilities and the relative expenditure for them at the household level. More than 46% of the total utility expenditure is for telephone services and the other 34% for fuel.

Table 2.7: Classification of Frequent Non-Food Expenditure Items into Regular and Non-Regular

Item	Code	Mean: 12m	Mean: 30d	Median: 12m	Median: 30d	Group
Kerosene oil	212	989	1,081	700	780	Regular
Coal, charcoal	213	2,014	1,951	500	0	
Cylinder gas (LPG)	214	7,975	7,708	7,200	2,880	
Matches, candles, etc	215	356	408	180	180	Regular
Ready-made clothing	221	4,684	6,761	3,000	0	
Cloth, wool, yarn, threads	222	3,060	3,700	2,000	0	
Tailoring expenses	223	992	1,160	700	0	
Footwear	224	1,828	2,710	1,200	720	
Toilet soap	225	619	630	480	480	Regular
Toothpaste, toothbrushes etc	226	451	434	300	300	Regular
Other personal care items	227	936	1,095	500	300	Regular
Dry cleaning, washing expenses	228	2,772	2,998	700	600	Regular
Personal services	229	531	493	300	240	Regular
Public transportation	231	2,969	4,194	1,500	1,200	Regular
Petrol, diesel, motor oil	232	14,681	16,263	8,400	9,600	Regular
Entertainment	233	809	1,013	400	240	
Newspapers, books, supplies	234	1,264	1,259	600	480	Regular
Pocket money to children	235	3,218	3,503	1,500	1,680	Regular
Wages paid to watchmen, etc	239	16,509	17,442	950	600	
Light bulbs, shades, batteries, etc	241	522	754	280	0	
Household cleaning article	242	1,147	1,204	800	840	Regular

Source: CBS (2011, p.28)
Note: Unweighted means and medians, including panel data.

Table 2.8: Utility Expenditure Summary (HH level)

Variable	Observations (HHs)	Mean (NRs.)	Std. Dev.	Share of Total Utility Expenditure	Minimum Value	Maximum Value
Garbage	5,988	103.05	442.14	0.96	0	8400
Electric	5,988	2109.21	4219.49	19.58	0	96000
Telephone	5,988	4948.39	10929.71	45.95	0	400000
Fuel	5,988	3608.96	5119.94	33.51	0	36000
Total Utility Expenditure	**5,988**	**10769.59**	**16942.96**	**100.00**	**0**	**447280**

Source: Author's calculations.

The inclusion of **health and education expenditures** in the consumption aggregate is also heavily debated in the literature. In general, education expenditures should not be included. This is mainly because they are irregular expenditures with temporal constraints, and the data used in this

dissertation is purely cross-sectional. For instance, one HH may have one child in school at the time the survey is conducted, while another child from a different HH may have finished school in the previous year. This would make the family with the child in school appear to have a higher consumption than the other family, introducing bias into the analysis. Deaton and Zaidi (2002, p. 33) argue that education expenditure should be smoothed over a lifetime or be excluded from the consumption aggregate. Also, education is seen as an investment in the future, rather than present-day consumption. It is thus argued that education should be included in the savings component of the monetary flow aggregate rather than in the consumption aggregate.[3] Moreover, the more children a HH has, the higher its education expenditures tend to be, consequently upward biasing the consumption estimate, despite families having to spend a larger proportion of disposable revenue on education with the trade-off of being able to consume less of other things. Despite all these reasons, the CBS includes education expenditure in their consumption aggregate.

Education-related expenses were surveyed in two sections of the survey. Once in section 7 relating to education expenses and once in section 6 as part of the non-food durable frequent goods category. In section 6, respondents could reply with a single number estimate for their educational expenses, while section 7 includes a much more complete list of education-related details (e.g. various school expenses, schooling hours, education degree attained). The CBS compared these two education expenses and mechanically chose the higher value of the two. This dissertation, however, took a two-way approach to integrating education expenses into the overall consumption aggregate. 1) To see whether education expenses should be included, education elasticities are calculated. The higher the elasticities, the greater the argument for inclusion according to Deaton and Zaidi (2002, p.34). Figure 2.3 illustrates the education (and health) elasticities. 2) The education expenditures for both sections are added separately and then individually compared for extreme values. **Conditional** marginal effects using linear predictions. Per capita consumption (*pccons)* serving as dependent variable.

Table 2.9 shows the results for the **predicted marginal elasticities for education (and health)** on consumption expenditure. Figure 2.3 provides the corresponding graphical representations. The elasticities are positive and increasing in both cases. This suggests that educational expenditure is judged as important for all households and should thus be included in the consumption aggregate. The elasticity estimates are also relatively high compared to some of the empirical estimates given by Deaton and Zaidi (2002, p. 33).

Since the educational (and health) elasticities are judged to be relatively high, it was decided to include educational (and health) expenses into the consumption aggregate. However, this dissertation does not use the single education expense estimate from section 6 due to the estimation error reasons mentioned earlier in this section, but the more refined estimates from section 7 of the survey. These estimates include those for tuition fees, other fees (exams, admission, events, etc), uniform, textbooks and supplies, transportation, private tuition, others (snacks, tea, etc.), and scholarships received over the past 12 months.

It is important to note, that Nepal has a dual education system, public and private. Given that the public education system lacks in quality (but is almost free of charge), families try to send their children to private schools (the choice spectrum of private schools is extensive, both in terms of pricing and quality, even in many remote rural areas). This means that the poorest spend relatively little on education (even in relative terms to their overall consumption), while those who can afford it, once a minimum disposable income threshold is passed, send their children to private schools, disproportionately increasing the share of educational consumption. Yet, the per capita education

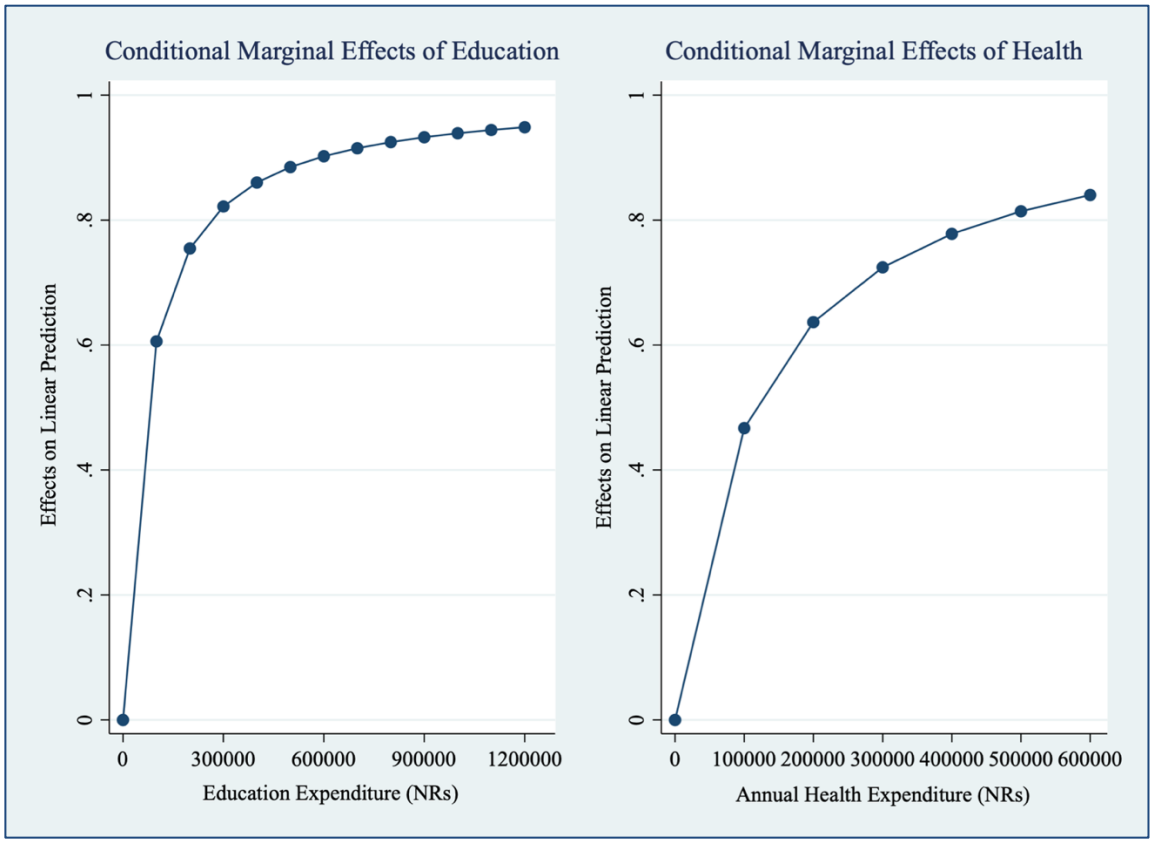
[3] Savings are aggregate components for many economic welfare indicators. However, they do not figure in income aggregates generally. Also, savings is not included in the income aggregate in this dissertation.

consumption share of overall consumption is 6.5% (Table 2.21). It is important to note, that due to this dual education system, wealthier HHs have a higher proportion of educational expenses than lower income HHs.

Just like educational expenses, the inclusion of **health care** expenses into the consumption aggregate is heavily debated in the literature. This is mainly due to measurement challenges, variations in terms of pricing (insurance, public, and/or private) and variance in access to health care treatment and medication, etc. Furthermore, one main normative argument suggests that when someone is sick, the situation is considered as a loss of welfare, and any expense related to health care or medication purely serves to regain the previous level of welfare. Additionally, it is difficult to then quantify this loss in welfare. Adding medical expenses would thus overestimate the aggregate. Also, health care expenses tend have a lumpy and irregular nature, and it is difficult to distinguish between "necessary" and "unnecessary" health expenses, e.g. vitamin tablets. Last, insurance coverage varies across the population, and those who have coverage either have private, public, or public and private insurance. So, in general, it is argued that health expenses should be excluded from the consumption aggregate, as is done by the CBS and Furuta (2016). In this dissertation, however, health care related expenses are included in the consumption aggregates. Just like the argument for education elasticity, Deaton and Zaidi (2002) suggest that the higher the consumption elasticity for health care, the stronger the argument for inclusion (see Figure 2.3 and Table 2.10).Conditional marginal effects using linear predictions. Per capita consumption (*pccons)* serving as dependent variable.

Table 2.9 indicates relatively high predicted health care elasticities, which are also relatively high compared to other empirical results given by Deaton and Zaidi (2002, p. 33). Health care expenses that are included in the consumption estimate comprise modern medicines and health services (e.g. doctor fees and hospital charges) and various traditional medicines and health services.

Figure 2.3: Education and Health Elasticities



Source: Author's calculations.

Note: Different x-axis scales: education maximum: 1,2 million NRs. and health max: 600,000 NRs. Conditional marginal effects using linear predictions. Per capita consumption (*pccons)* serving as dependent variable.

Table 2.9: Predicted Marginal Elasticities and Overall Average Elasticity for Education and Health on Consumption Expenditure

Predicted Marginal Elasticities for Health on Per Capita Consumption Expenditure						
Education	**ey/ex**	**Std. Err.**	**t**	**P > I t I**	**[95% Conf. Interval]**	
1	0	-	-	-	-	-
2	0.6058	0.0094	64.18	0.0000	0.5873	0.6243
3	0.7545	0.0073	103.06	0.0000	0.7402	0.7689
4	0.8218	0.0058	141.94	0.0000	0.8104	0.8331
5	0.8601	0.0048	180.82	0.0000	0.8508	0.8694
6	0.8849	0.0040	219.7	0.0000	0.8770	0.8928
7	0.9022	0.0035	258.59	0.0000	0.8953	0.9090
8	0.9150	0.0031	297.47	0.0000	0.9089	0.9210
9	0.9248	0.0027	336.35	0.0000	0.9194	0.9302
10	0.9326	0.0025	375.23	0.0000	0.9277	0.9375
11	0.9389	0.0023	414.11	0.0000	0.9345	0.9434
12	0.9442	0.0021	452.99	0.0000	0.9401	0.9482
13	0.9486	0.0019	491.87	0.0000	0.9448	0.9523
Predicted Marginal Elasticities for Health on Per Capita Consumption Expenditure						
Health	**ey/ex**	**Std. Err.**	**t**	**P > I t I**	**[95% Conf. Interval]**	
1	0	-	-	-	-	-
2	0.4669	0.0155	30.17	0.0000	0.4366	0.4973
3	0.6366	0.0144	44.26	0.0000	0.6084	0.6648
4	0.7243	0.0124	58.35	0.0000	0.7000	0.7487
5	0.7780	0.0107	72.44	0.0000	0.7569	0.7990
6	0.8141	0.0094	86.53	0.0000	0.7957	0.8326
7	0.8401	0.0084	100.62	0.0000	0.8238	0.8565
Overall Average Education and Health Elasticities on Per Capita Consumption Expenditure						
	ey/ex	**Std. Err.**	**t**	**P > I t I**	**[95% Conf. Interval]**	
Education	0.1352638	0.0036097	37.47	0	0.1281874	0.1423402
Health	0.063571	0.0031349	20.28	0	0.0574255	0.0697164

Source: Author's computations.

Note: Conditional marginal effects using linear predictions. ey/ex represents the computed elasticity. Per capita consumption (*pccons)* serving as dependent variable.

The third and final consumption expenditure aggregate component is **housing**. Table 2.21 shows that an average of over 18% of overall consumption expenditure is dedicated to housing, making this an important component of welfare. However, calculating housing expenditure in reality is problematic. Deaton and Zaidi (2002, p. 36–38) conclude that for extreme cases, the housing component should be excluded. There are three main reasons for this. First, if real estate is bought in the 12 months prior to taking the survey, it should not be included since it is a rare and large expense.

Second, most LSMS questionnaires ask respondents to estimate the rent of their accommodation in case they do not pay rent or own the dwelling. This hypothetical concept is problematic because often the interviewees do not have sufficient market knowledge to correctly estimate the value of their housing, and often the responses are not credible. Respondents tend to have an incentive to underreport the value of their assets.

Last, in many countries and particularly in the developing country context, real estate markets are not sufficiently established to allow for serious estimates. When **rental equivalents** are potentially inaccurate, the benefits of completeness need to be weighed against the costs of error.

Despite all these reasons, an attempt is made to appropriately including housing into the consumption aggregate following the Deaton and Zaidi (2002, p.77–81) criteria. When rental values for dwellings are given, these are then annualised and integrated into the overall consumption aggregate. Furthermore, when the dwelling is owned by the household or received free of charge, an estimate of the annual rental equivalent is included in the consumption aggregate. If the rental value is estimated by the respondent (e.g. because he is living in subsidised housing or the with family), then the data is inspected prior to use. Some unreasonable rent prices are reported (i.e. less than 100NRs./month or above 30,000NRs./month according to the CBS (2011)). Some observations are below 100NRs./month, yet they were deemed as acceptable by the author, judging by other household and per capita expenses. Also, the author set the upper limit to 250,000NRs./month after exchanges with Nepalese contacts and employees from the CBS. Some observations are drastically large. The CBS drops over 68 observations and the author 6.

Furthermore, when there are missing observations, a **hedonic housing regression** is developed to estimate the rental values. This is a common procedure. Numerous elements that contribute to the quality of the dwelling are included, such as roofing, type of sanitation of wall material, access to facilities (e.g. paved roads), and some regional dummies. Their measurement allows for rental parameter estimations. These then serve to compute the rental values for the missing observations. When respondents' rental estimates appear unrealistic, then the hedonic housing regression also replaces these values. Table 2.10 reports the estimated hedonic housing regression coefficients by the author and by the CBS (2011). Some values are similar, while some are different. This is due to the utilisation of different approaches. This dissertation focuses on the approach used by Deaton and Zaidi (2002), which further includes geographical elements, such as housing being in the capital of Kathmandu, in other urban areas, or in the rural areas of the hill region. These estimates are then used to predict an annual rental equivalent for missing or implausible observations.

Table 2.10: Hedonic Housing Regression Coefficient Estimates by the Author and by CBS (2011)

	Coefficients, (Scholl)°	**Coefficients, CBS (2011)°°**
Kathmandu	0.730***	-
	(0.06)	-
Other urban areas	0.437***	-
	(0.04)	-
Hill region rural	-	-
	-	-
Terai region rural	0.096*	-
	(0.04)	-
Log(number of rooms)	0.406***	0.103***
	(0.03)	(0.01)
Log(total area of dwelling)	0.144***	0.179***
	(0.02)	(0.02)
Log(value of durable goods)	0.254***	0.060***
	(0.02)	(0.01)
Dwelling has a kitchen	0.231***	0.268***
	(0.03)	(0.02)
Paved road next to dwelling	0.154*	0.432***
	(0.07)	(0.03)
Dwelling has cemented wall	0.118***	0.269***
	(0.03)	(0.04)
Dwelling has cemented foundation	-0.350***	0.241***
	(0.04)	(0.04)
Cemented or tin roof	0.211***	0.214***
	(0.03)	(0.03)
Dwelling has a window	0.304***	0.210***
	(0.03)	(0.03)
Piped water supply	0.037	-0.089***
	(0.03)	(0.03)
Communal garbage collection	0.437***	0.207***
	(0.05)	(0.04)
Piped water inside dwelling	0.271***	0.201***
	(0.06)	(0.03)
Electricity for lighting	0.183***	0.183***
	(0.04)	(0.03)
Dwelling has telephone facility	0.468***	0.359***
	(0.04)	(0.03)
constant	4.350***	4.198***
	(0.14)	(0.13)
Number of observations	5,262	6,861
R2	0.625	0.668
Adjusted R2	0.624	0.666

Note: * p<0.05, **p<0.01, *** p<0.001. Standard errors are in parenthesis.
° Hedonic regression estimates by the author. °° Hedonic housing regression estimates by CBS (2011).

In addition to the suggestions by Deaton and Zaidi (2002), some goods that fall under the category of **own account production and consumption** were added. These include, for example, rope, firewood, tailoring, shoemaking, water fencing, warehouse repairing, biogas, and pickle. The questionnaire only asks according to the typical month criterion and for annual estimates. These are compared and checked for outliers. The monetary value of the good at the local market price for the annual estimate is the preferred value.

All the components for the consumption expenditure aggregate are summed to give the total HH consumption expenditure level per HH. This is then divided by the size of each HH to give the per capita consumption expenditure variable.

Table 2.11 summarises some descriptive statistics for the consumption expenditure aggregates following the recommendations by Deaton and Zaidi (2002), Furuta (2016), and the CBS (2011). The table shows both total HH and per capita consumption expenditures. Depending on which methodology is used (refer to Table 5.2 to see a summary comparison of the different methodologies), the consumption aggregates vary from as low as 306USD to 580USD per capita annual consumption expenditure. This highlights the fact that great care must be taken when cleaning the data and when computing the consumption aggregates. Depending on how consumption was aggregated, i.e. which outliers were dropped, what parts of the sample and how items were weighted, and so on, the value of the aggregate changes, and thus does the distribution of consumption across and within households. This in turn affects the following empirical analysis i.e. the inequality in outcome and the inequality of opportunity analysis. More on this in section 2.6, which compares and comments on the income and consumption aggregates' distribution, disparity, and inequality.

In this study, the variable *pccons* serves as the main per capita consumption expenditure dependent variable. It is judged to be the ideal aggregate by the author since it combines and weights the suggestions of various aggregation guidelines, taking into consideration the dataset characteristics. Table 5.2 is a summary of the guidelines followed and the adjusted aggregation methodology by the author. The following are some of the reasons as to why the variable *pccons* is regarded as the optimal aggregate. First, *pccons* uses the widely acknowledged typical monthly criterion for the annualised computation process. Second, at every stage of the aggregation process, all component items, units, and outliers, etc. are checked for consistency and reliability and were adequately dealt with. This is opposed to particularly the CBS' mechanical procedure to eliminate consistently the bottom and top 1% of the distribution. Third, *pccons* includes both education and health consumption, as was suggested by the respective elasticity analyses.

Pccons2 (and *pcconsWEEK2*) follow the same aggregation methodology as *pccons* and *pcconsWEEK*, respectively, but excluding both health and education consumption. This is done for comparison purposes.

The "WEEK" prefixes of the variables indicate that the "past week" criterion was used for comparison purposes whenever the information was provided. As expected, when the sole difference in aggregation methodology lies in the typical month and past week criteria, then the WEEK estimates are consistently smaller. For instance, when the Deaton and Zaidi's (2002) guidelines are followed and the only difference is this aggregation criterion, then per capita consumption expenditure is estimated at over 580USD/annum using the typical month and 398USD/annum using the past week criterion.

PcconsNLSS is the aggregate following the CBS (2011) aggregation methodology guidelines, one of the most impacting methodological differences being the consequent mechanical elimination of the top and bottom 1% of the sample.

Furuta (2016) advises to exclude education and does not speak of health expenditures in his aggregation methodology. Variable *pcconsFURUTA* furthermore calculates the durable goods depreciation rates in a different manner.

Overall, this section briefly elaborated on the discrepancy between different consumption aggregation methodologies and elaborated on the methodology followed for the per capita consumption aggregate (*pccons*). The computed values and some descriptive statistics are given for the various consumption expenditure aggregates and *pccons* is seen as the adjusted per capita consumption expenditure aggregate. It serves as the key per capita consumption expenditure dependent variable for the inequality in output and the subsequent inequality of opportunity analyses. The descriptive statistics of the consumption aggregate and its component are discussed further in section 2.6.

Table 2.11: Consumption Expenditure Aggregates Based on Four Main Sets of Guidelines

	Variable name	Obs.	Mean (NRs.)	Mean (USD**)	Min.	Max.	Standard Deviation	Median	Interquartile Range	Coefficient of Variation
Author,s "optimized" consumption aggregation combining the recommendations from Deaton and Zaidi (2002), CBS (2011) and Furuta (2016)										
Total household consumption expenditure, typical month criteria, Scholl aggregation methodology	cons	5,987	215,029.00	2,691.23	9,056.00	4,271,874.00	219,449.70	152,497.30	141,634.40	1.02
Per capita consumption expenditure, typical month criteria, Scholl aggregation methodology	pccons	34,063	**44,091.71**	**551.84**	5,689.14	741,725.60	46,537.52	30,111.22	29,891.09	1.06
Total household consumption expenditure, weekly criteria, Scholl aggregation methodology	consWEEK	5,987	145,492.70	1,820.93	7,198.67	4,104,585.00	193,046.20	87,224.28	103,059.10	1.33
Per capita consumption expenditure, weekly criteria, Scholl aggregation methodology	pcconsWEEK	34,063	**29,570.03**	**370.09**	2,777.31	721,115.10	40,762.86	17,018.93	20,903.36	1.38
Total household consumption excluding education & health	cons2	5,987	190,156.00	2,379.92	9,056.00	3,816,014.00	197,180.60	136,621.30	119,311.60	1.04
Per-capita annual consumption excluding education & health	pccons2	34,063	**39,015.90**	**488.31**	4,841.75	720,363.80	41,936.26	26,800.36	25,612.98	1.07
Total household consumption excluding education using past week criteria & health	consWEEK2	5,987	120,619.70	1,509.63	6,798.67	3,648,725.00	171,056.80	71,838.28	77,873.09	1.42
Per-capita annual consumption excluding education using past week criteria & health	pcconsWEEK2	34,063	**24,494.23**	**306.56**	2,216.20	663,676.30	36,174.79	13,830.93	16,194.96	1.48
Following the Deaton and Zaidi (2002) guidelines										
Total household consumption expenditure, Deaton and Zaidi (2002) aggregation methodology	consDZ	5,987	227,849.90	2,851.69	7,056.00	48,300,000.00	975,349.30	145,060.70	137,476.60	4.28
Per capita consumption expenditure, Deaton and Zaidi (2002) aggregation methodology	pcconsDZ	34,063	**46,350.95**	**580.11**	4,497.89	30,000,000.00	265,192.50	28,404.29	28,379.85	5.72
Total household consumption expenditure, Deaton and Zaidi (2002) aggregation methodology, weekly criteria	consWEEKDZ	5,987	158,313.70	1,981.40	6,598.67	48,200,000.00	969,203.10	79,817.79	99,536.44	6.12
Per capita consumption expenditure, Deaton and Zaidi (2002) aggregation methodology, weekly criteria	pcconsWEEKDZ	34,063	**31,829.28**	**398.36**	1,989.54	30,000,000.00	264,026.70	15,362.63	19,709.59	8.30
Following the CBS (2011) guidelines										
Total household consumption expenditure, NLSS III Report aggregation methodology	consNLSS	5,987	184,562.60	2,309.92	7,056.00	2,387,767.00	163,316.50	135,581.40	127,884.50	0.88
Per capita consumption expenditure, NLSS III Report aggregation methodology	pcconsNLSS	34,141	**37,879.21**	**474.08**	4,280.58	548,502.20	35,023.22	26,553.68	26,815.59	0.92
Following the Furuta (2016) guidelines										
Total household consumption expenditure, Furuta (2016) aggregation methodology	consFURUTA	5,987	216,455.80	2,709.08	7,056.00	4.83E+07	974,404.50	135,662.00	129,049.40	4.50
Per capita consumption expenditure, Furuta (2016) aggregation methodology	pcconsFURUTA	34,141	**44,022.19**	**550.97**	4,276.58	3.00E+07	264,711.00	26,549.68	26,998.07	6.01

*Where *totcons* is total household consumption and *pcpcons* is per-capita annual consumption. Both utilise the "typical month" criterion. *totconsWEEK* is the total household consumption utilising the "past week" criterion while *pcpconsWEEK* is the per-capita annual consumption using the "past week" criterion.

** Average exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

2.4.2. Robustness Checks for the Consumption Expenditure Aggregate

There are countless ways to check the robustness at each stage of the aggregation process and of the final consumption aggregate. Deaton and Zaidi (2002) elaborate extensively on how to do so. They classify the approach to checking the consumption aggregate robustness into three main categories.

1) Adjusting for differences in the cost of living across the sample. This includes procedures such as calculating the Paasche Price Index and calculating the Laspeyers indices.
2) Adjusting for HH composition. This includes, for instance, equivalence scale measures; a more behavioural, more subjective, and purely arbitrary approach, as well as finally adjusting the consumption aggregate for both differences in living costs and household compositions.
3) Checking for sensitivity. These include, for instance, the stochastic dominance analysis and the estimation of the effect of measurement error.

This dissertation mainly focuses on the analysis of inequality in outcome and inequality of opportunity using the NLSS III dataset. It is thus outside the scope of this dissertation to carry out all these robustness checks for the consumption aggregate. However, during the data cleaning and the aggregation process, great care was taken to carefully weigh the inclusion and exclusion of certain elements and determine how to best deal with outliers and how to optimise the aggregation methodology to suit the NLSS III dataset, and so forth (see elaboration above). Lastly, a descriptive comparison is also done between the final aggregates.

2.4.3. Limitations of the Consumption Expenditure Aggregate and Further Extensions

There are a series of **limitations** in the aggregation process, many of which are related to the dataset itself. The NLSS III dataset excludes all institutional (e.g. people living in prisons, school hostels, or hospitals) and diplomatic HHs. However, foreigners who officially reside in Nepal are included.

Moreover, much of the data is based on the recollection memory of the respondents. This is prone to great recollection error and recall bias. The CBS, however, assures that maximum efforts were made to train the interviewers so that they could obtain the most reliable information from the respondents. Also, the CBS guarantees that a strong emphasis was placed on the data collection teams to perform consistency checks of the data upon entry (e.g. missing or implausible data) and when initially cleaning the raw data on the ground.

The length of the survey is another cause for error during the data collection process. There is always a trade-off between the inclusion of a maximum number of questions and the quality of the data provided. The CBS assures that maximum efforts were taken to alleviate the burden on interviewees and, for instance, that data were automatically entered by the software or the interviewer where possible.

Furthermore, local measurement unites were used. Despite the provision of a conversion table, it is easy to make measurement errors. For instance, quantities are measured in over 10 different units. The most conventional units are kilograms, litres, or pieces, but there are also numerous indigenous measurements, such as muri, pathi, manna, or kuruwa. However, during the entire data cleaning and aggregation process, units were cross-checked and verified by the CBS and again by the author.

The surveyed sample population is based on the sampling methodology of the NLSS I and II rounds. These were initially based on the national labour force surveys. Thus, the author was told by employees of the CBS that emphasis was placed on gathering nationally representative socio-economic information, while household composition was not a priority. Thus, for instance, the sex ratio in the

NLSS III is below the national average as estimated by the population census. More on this data shortcoming in section 3.3.2.3.

As previously explained, there is no consensus on how to aggregate economic outcome. Thus, this study develops an "optimised" aggregation methodology adapted to the NLSS III dataset following various guidelines. This, however, makes the international comparison of the aggregates difficult.

There are a number of ways through which the **data quality and aggregation process could be improved.** The NLSS III dataset, being a digitally entered survey, should allow for much more reliable data entry and minimal human error. Yet, there are some dataset imperfections. A richer, more detailed, and larger dataset could help to improve the representativeness of the dataset and add in terms of variable completion (e.g. more details on and more consumption items in the survey). Also, not all sections bear both typical month and past week information. This could provide additional valuable information. However, there is always a trade-off between collecting maximum information and receiving quality inputs from the respondents. Deaton (1997) gives a more complete picture on how to design, execute, and analyse HH surveys in order to allow researchers to maximise the output of the datasets.

2.5. Income Aggregation

This section first elaborates on the methodologies used to compute the income aggregate using the NLSS dataset. It then highlights various robustness checks. Finally, this section states some limitations and further improvement possibilities for aggregating income.

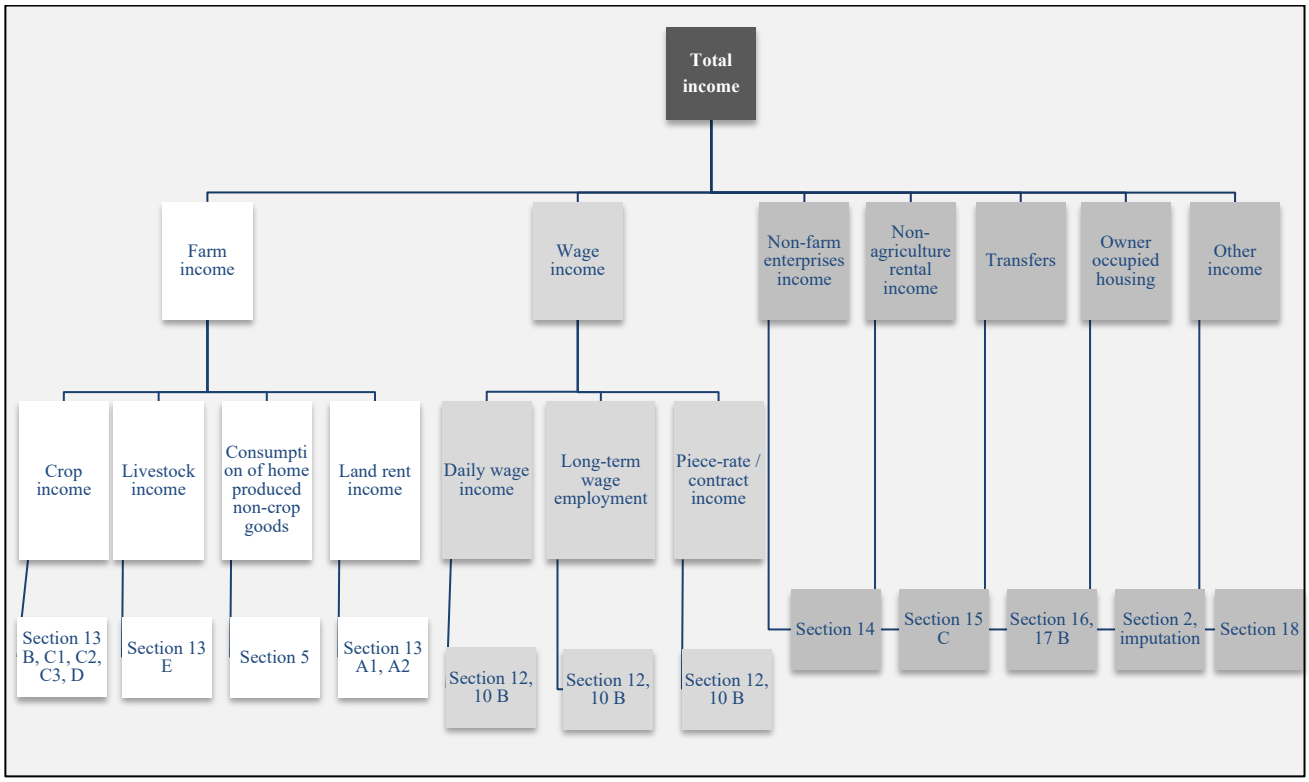
2.5.1. Methodology: Guidelines for Aggregating Income Using the NLSS Dataset and Some Descriptive Statistics

The aggregation methodology for income is slightly more straightforward than the one for consumption expenditure. In principle, income represents the flow of monetary means, and it is therefore more accurately captured by the dataset. However, in the developing country context, few people actually receive wages in terms of monetary transfers. So, other components are also included in the aggregate. Figure 2.4 illustrates the main relevant components and sub-components (and relevant survey sections) of the overall income aggregate. The illustration is based on the computation methodology by the CBS (2011, Ch. 11). This dissertation mainly follows the CBS' recommendations with minor adjustments. Other main references for the income aggregation methodology include Coulombe and McKay (2008), Food and Agriculture Organization (2008), Quinones et al. (2009), International Labour Organization (2003) and Furuta (2016). A summary for the component elements on the inclusion, exclusion, or adjustment of various items can be found in Table 2.12.

The main income components are:

1) Farm income
2) Wage income
3) Non-farm enterprises income
4) Non-agricultural rental income
5) Transfers
6) Owner occupied housing
7) Other income

Figure 2.4: Components of the Income Aggregate (Main Components, Sub-Components, Relevant Sections in Survey)



Source: Author's illustration based on CBS (2011, Vol. 2, Ch. 11)

Farm income comprises four main sub-components, namely crop income, livestock income, land rent income, and the consumption of home-produced non-crop goods. **Crop income** is one of the most challenging components to add to the income aggregate. Own-produced crops can be sold at the market, home-consumed, traded, or used as in-kind payment for obtained services and goods. However, an attempt is made to conduct a gross price imputation for all harvested crops (both CBS (2011) and the author do this). Often, the unit selling price is unavailable. In this case, they are then computed using a progressively increasing average price at four different levels (ward, district, region, and country). Whenever prices are missing, the prices are imputed, and implausible cases are dropped. Furthermore, some HH use in-kind payment to their landlords for the use of land. This is deducted from the total harvest volume. The harvest values are therefore the net value of in-kind rent payments to the landlord.

Further, **agricultural earnings** (e.g. agricultural by-products, renting out animals, tractors etc) are added to crop income. Agricultural expenditures (e.g. fertilizers, hired labour, water) are deducted from crop income.

Income from **livestock** is the value of the net sale and purchase of these and their by-products.

The **consumption of home-produced non-crop** goods includes, for instance, eggs, buffalo meat, or chickens and are added. **Tobacco** is taken out following the guidelines of Deaton and Zaidi (2002).

Land rent income includes both net monetary and in-kind income. This, however, excludes the previously deducted in-kind payments to landlords.

Wage income has three sub-categories, namely daily wage income, long-term wage income, and piece-rate and contract income. Daily wage income includes daily cash and any in-kind payment received for labour work. This is then multiplied by the amount of days worked.

Long-term wage income is divided into agricultural wages received and other sector wages. Agricultural wages also include both daily cash and in-kind payments, which are multiplied by the number of days worked. Wages received for work other than agricultural work are computed on a monthly basis and further include elements such as bonuses, tips, and allowances.

Piece-rate and contract-based income is computed either by piece or per project in the period of the surveyed year.

Wage income is aggregated at the HH level and added to the overall income aggregate.

Non-farm enterprise income is the net revenue of non-farm enterprise sales minus the related expenditures for raw materials (e.g. wages, in-kind payments, and other operating expenses). This is also done at the HH level.

Non-agricultural rental income includes any property that is rented out by the HH.

Transfer income includes any cash or in-kind transfers received by the HH, particularly through remittances.

Owner-occupied housing income follows the same methodology as iterated in the housing consumption section for the consumption aggregate. Refer to Section 2.4.1 for more details.

Other income sources include all cash and in-kind received revenue through savings, shares and stocks, social assistance, and pensions, etc. Note, it is not the overall savings value that is included, just the revenue arising from them (e.g. through interest rates).

Some **income expenses**, such as income, wealth, and property tax are deducted from the overall income aggregate. When these expenses are deducted, the remaining income is referred to as disposable income. The variable *pcincome* in Table 2.13 is the disposable income, and *pcincome1* is the pre-tax income level. The CBS (2011) report on the NLSS III does not refer to tax deduction in its income aggregate, and neither does Furuta (2016). Quinones et al. (2009, Section 3.3.2.), however, suggest that expenses such as social services and other taxes, should be deducted from total income, the aim being to create a net income aggregate, i.e. disposable income. The ILO (2003) also suggests using disposable rather than total earned income.

The NLSS III dataset does not provide much information on **direct and indirect taxes** paid by households and individuals. Paid taxes only figure in two parts of the NLSS survey, and with little detail. They appear 1) in the long-term non-agriculture wage section, where the individuals are asked only whether they pay taxes, but not how much; and 2) in non-food infrequent and durable goods expenses, where only two questions are reserved for taxes: a) whether taxes (income tax, property, and wealth taxes) were paid, and b) if so, how much over the past year in total. The latter tax expenses were then deducted from overall income in order to provide a disposable income aggregate (*pcincome*), which serves as the main reference variable. It is important to note that when taxes are paid, the mean level of paid taxes is only 402.67NRs. per capita (standard deviation of 4,213 and maximum value of 250,000 NRs.). This minor tax deduction makes virtually no difference to the mean income aggregate. It is an average of ca. 90NRs. less per capita for the 34,146 individuals.

Some items are omitted due to their particular character. For instance, property sales are excluded due to their bulky and infrequent nature. Likewise, property repair and maintenance fees are not included since they characterise investment and not income flows.

All of the above components are added to give an overall HH-level income aggregate. The per capita level is then calculated. These income aggregates are pre-tax. Table 2.13 summarises the income

aggregates at the household and the per capita levels following different aggregation methodologies. It also shows some descriptive statistics. Depending on the aggregation methodology, per capita income ranges between 608USD and 1,247USD per annum. Just as for the consumption aggregate, this highlights the fact that great care must be taken when cleaning the data and when computing the income aggregates. Depending on how it is aggregated, i.e. which outliers were dropped, what parts of the sample and how items were weighted, and so on, the value of the aggregate changes, and thus so does the distribution of income across households. This in turn then affects the empirical analysis that follows, i.e. the inequality in outcome and the inequality of opportunity analysis. More on this in section 2.6 which compares and comments on the income and consumption aggregates, distribution, disparity, and inequality in outcome indices.

In this study, the variable *pcincome* will serve as the main per capita income dependent variable. It reflects disposable income, since taxes (income tax, wealth, and property tax) are deducted from the overall "ideal" income aggregate, *pcincome1*. *Pcincome1* is the pre-tax per capita income. *Pcincome* is judged by the author to be the "ideal" income aggregate following the "optimised" aggregation methodology developed by the author given the literature and the nature of the NLSS dataset.

Pcincome2 stands for per capita income excluding utilities, following the advice of the literature to exclude some utilities due to the divergence in geographical coverage of utility provision, subsidization, and quality. Since the dataset provides relatively high-quality information on utilities, only water is excluded (a similar methodology as for the consumption aggregate).

Pcincome3 is the *pcincome1* estimate excluding wage income. This is used because the CBS (2011) does their socio-economic analysis for a) total per capita income, b) for per capita income excluding wage income, and c) per capita wage income. So as a reference, Table 2.13 also provides the pcincome3 aggregate.

Pcincome4 is the per capita income estimate following the CBS (2011) methodology on mechanically eliminating the top and bottom 1% of the distribution of the sample. The CBS considers these as the outlier cut-off lines. This procedure is judged to be too mechanical by the author. As for the consumption aggregate, each component and item for the income aggregate at every level of the aggregation process is checked for outliers, possible unit measurement errors, missing values, or clearly implausible values.

Overall, this section showed that there are various methodologies to compute the income aggregate, and the results are shown in Table 2.13. A summary of items to include and exclude are found in Table 2.12. The computed values and some descriptive statistics are given for the various income aggregates and *pcincome* is seen as the fine-tuned per capita income aggregate. It serves as the key per capita income dependent variable for the inequality in outcome and the subsequent inequality of opportunity analyses. There is more on the descriptive statistics of the income aggregate and its component in section 2.6.

Table 2.12: Details for the Income Aggregate Components

Main components	Items to add	Items to deduct
Farm income	+ Value of total crop production (net of share paid to landlord) + Value of by-product production + Net income from renting farm assets (draft animal, tractor, thresher etc.) + Value of sales from non-crop farm production (milk, ghee, eggs, etc.) + Earnings from the sale of livestock - Expenditure for the purchase of livestock + Value of home-produced non-crop consumption	- Cultivation costs (seeds, fertilizers, hired labour, irrigation etc.) - Maintenance expenditures on farm machinery and buildings - Fodder and other livestock expenditure (veterinary services)
	+ Total cash and in-kind received from tenants on land leased-out	- Cash rent paid to landlord on leased land
Wage income	+ Value of cash and in-kind earning per year in agriculture (includes daily, piece-rate, and permanent labour) + Value of cash and in-kind earning per year outside agriculture (includes daily, piece-rate, and permanent labour)	
Non-farm enterprises income	+ Gross income from non-agriculture enterprises/activities during the past 12 months	- Wage paid both cash and in-kind - Energy expenditure - Expenditure on raw material - Other operating expenditure - Share of net revenues paid to partners (or kept by partners)
Non-agriculture rental income	+ Income from renting out non-agriculture property + Income from renting out non-agriculture assets	
Transfers	+ Remittances (cash and in-kind)	
Owner-occupied housing	+ Imputed rental value of housing which would have had to be paid to purchase housing services	
Other income	+ Interest, dividends, profit earning from shares, and savings/deposit accounts + Social security payment + Pension income (domestic and foreign) + Commission fees and royalties, other incomes	- Income expenses, such as income, wealth, and property taxes

Source: Author's summary of basic guidelines from CBS (2011, Ch. 11), with some adjustments.
Note: All components are checked for outliers before aggregation. On crop income: tobacco consumption is taken out. For details see Deaton and Zaidi (2002).

Table 2.13: Income Aggregates*

	Variable name	Obs.	Mean (NRs.)	Mean (USD*)	Min.	Max.	Std. Dev.	Median	Interquar-tile Range
Household disposable income	totincome	5,988	461,520.2		-5592454.00	18400000.00	754732.8	252832	414710
Per capita disposable income	**pcincome**	**34,146**	**97,028.35**	**1,214.37**	**-2,796,227.00**	**4,437,499.00**	**150,070.40**	**56,114.50**	**90,671.00**
Household income	totincome1	5,988	461,954.10	5,781.65	-5,592,454.00	18,400,000.00	755,142.60	252,904.50	415,750.50
Per capita income	pcincome1	34,146	97,118.68	1,215.50	-2,796,227.00	4,437,537.00	150,185.00	56,147.00	90,784.41
Household income including utilities	totincome2	5,988	473,546.30	5,926.74	-5,539,594.00	18,400,000.00	770,234.80	263,587.00	429,687.00
Per capita income including utilities,	pcincome2	34,146	**99,669.48**	**1,247.43**	-2,769,797.00	4,439,589.00	153,882.50	57,209.17	91,329.67
Household income excluding wage income	totincome3	5,988	236,900.90	2,964.97	-5,592,454.00	7,417,012.00	388,088.00	132,377.00	180,782.40
Per capita income excluding wage income,	pcincome3	34,146	**48,614.26**	**608.44**	-2,796,227.00	2,517,734.00	84,716.56	26,003.75	36,311.80
Household income with CBS outlier range of 1-99% of sample	totincome4	5,870	428,961.20	5,368.73	5,335.00	6,995,602.00	523,811.20	259,246.50	404,009.00
Per capita income with CBS outlier range of 1-99% of sample	pcincome4	33,590	**89,692.54**	**1,122.56**	3,551.43	659,041.00	96,268.47	55,714.25	88,617.83

Note: * following various guidelines, on e.g. the inclusion or exclusion of utilities and wage income and on how to deal with outliers. ** average exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

2.5.2. Robustness Checks for the Income Aggregate

There are countless ways to check the robustness and consistency during the aggregation process and of the final income aggregates.

This dissertation mainly focuses on the analysis of inequality in outcome and inequality of opportunity using the NLSS III dataset. It is thus not within the scope of this dissertation to carry out many further robustness checks than what have already been done. During the data cleaning and the aggregation process, great care was taken to carefully weigh the inclusion and exclusion of certain elements and determine how to best deal with outliers and optimise the aggregation methodology to suit the NLSS III dataset, and so forth (see elaboration above).

2.5.3. Limitations of the Income Aggregate and Further Extensions

As for the consumption aggregate, there are a series of **limitations** in the aggregation process, and many of these are related to the dataset itself. Refer to Section 2.4.3 for more details on 1) the exclusion of certain HHs from the dataset, 2) the memory recollection bias of respondents and survey length, and 3) the measurement error for units.

Moreover, Deaton (1997) states that the notion of income, especially in developing countries, is problematic. This is because most HHs engage in agricultural, home-produced, or family business activities, meaning the notion of income appears abstract to the respondents when asked about it. So, most surveys impose an accounting framework when collecting and analysing the data. This was applied in the various aggregation processes.

Also, Deaton (1997) highlights the challenge respondents face to accurately evaluate their assets and the returns of their assets. Information is often incomplete, and respondents tend to underestimate their assets.

Furthermore, the NLSS III dataset does not provide much information on **direct and indirect taxes** paid by households and individuals. Paid taxes only figure in two parts of the NLSS survey, and in little detail. Further, one cannot conclude as to what and at which value public goods are rendered to individuals for their taxes. Thus, the aggregate does not fully reflect economic welfare. Given that

consumption expenditure measures individuals' valuation of consumption, it acts as an important comparison to the income aggregate.

As previously explained, there is no consensus on how to aggregate economic outcome. Thus, this study develops an "optimised" aggregation methodology adapted to the NLSS III dataset following various guidelines. This, however, makes the international comparison of the aggregates difficult.

There are a number of ways through which the **data quality and aggregation process could be improved.** A richer, more detailed, and larger dataset could help to improve the representativeness of the dataset and add in terms of accuracy. Section 2.4.3 on data limitations and possible improvement suggestions for the consumption aggregate gives more details on this. These are equally relevant for the income aggregate.

2.6. Descriptive Comparison of the Income and Consumption Expenditure Aggregates: Distribution, Disparity, and Inequality

The distribution, disparity, and inequality of income (or consumption) can be described in two main ways. First, at the descriptive level, where the distribution and disparities of the aggregates are highlighted using, for instance, graphical means, like the Lorenz Curve, or descriptive statistics that decompose the data, for example by population or geographical factors.

Second, at the statistical level, where inequality of outcome is computed using inequality indices that analyse the distributional characteristics of the economic outcome variables. Examples are the Gini Coefficient, the Theil Index or the Mean Logarithmic Deviation.

Section 2.6 first compares and describes the various per capita income and consumption aggregates. It then describes the distribution and disparity of each aggregate, using geographical, demographic, and basic distributional statistics. Last, this section compares a selection of inequality of outcome indices computed for the various aggregates.

2.6.1. Comparison Within and Between the Income and Consumption Expenditure Aggregates

Table 2.14 compares the per capita income and consumption aggregates computed following various aggregation guidelines, including the author's adjusted aggregation methodology for the NLSS III dataset. It furthermore provides aggregate information from the CBS (2011) reports, which also used the NLSS III datasets to compute their aggregates, and from the WDI (different data, 2011).

Referring to the author's computations, one can see that there are large variations between the overall per capita aggregates depending on the guidelines followed. For consumption, the aggregate can be as low as 370USD and as high as 580USD per capita per year depending on the aggregation methodology followed. The latter is almost double the smallest aggregate. As explained in Section 2.4.1, the estimates using the typical month criterion are consistently higher than those using the last week criterion. The proposed reference variable *pccons* by the author is computed at 552USD, which is 100USD over what the CBS (2011) had computed (436USD). This is due to some minor methodological differences and well as the significant impact of mechanically eliminating the bottom and top 1% of the distribution. In particular, dropping the top 1%, who hold a relatively large proportion of the national consumption level, has an important impact on the mean level of the

aggregate. The World Bank estimates for 2011 are at 463USD but using different dataset sources (World Bank national accounts data and OECD National Accounts data files).

For income, the aggregate can be as low as 608USD and as high as 1,247USD per capita per year depending on the aggregation methodology followed. The latter is also more than double the smallest aggregate. The proposed reference income variable by the author for the continued analysis is *pcincome*. It adds to over 1,214USD per capita per annum. This value is almost double the value of that of the CBS (2011; 521USD). This is again due to the fact that the CBS mechanically eliminated the bottom and top 1% of the distribution. Particularly the housing and wage components are affected by this. This lowers the estimate by almost 200USD. Furthermore, in informal exchanges with CBS staff, the author was told that, furthermore, taxes were deducted using confidential data from the tax authorities. This could not be obtained by the author and only the limited tax information included in the data was used for the *pcincome* aggregate. There are also other minor methodological differences, as well as the exclusion of wage income in most of the CBS' further socio-economic analysis.

The *pcincome* value is almost double that of the World Bank (WDI, 2011; 643USD). It is important to note that the World Bank used a different data source for their aggregation (World Bank national accounts data, and OECD National Accounts data files). Also, it is an adjusted net aggregate, deducting direct and indirect taxes. This significantly lowers their disposable income estimate compared to the one of the author, due to the lack of complete data provision on tax payments in the NLSS III dataset.

As previously explained (literature review in Section 2.1 on economic welfare measures), the income aggregates are consistently higher than those of consumption expenditure. Ferreira and Gignoux (2011) also highlight the robustness of this and the further importance this has when interpreting the econometric inequality of opportunity results.

Table 2.14: Comparison of Per Capita Income° and Consumption Expenditure Aggregates

Variable name	Obs.	mean (NRs.)	mean (USD*)
Author's "optimised" consumption aggregation methodology**			
pccons	**34,063**	**44,091.71**	**551.84**
pcconsWEEK	34,063	29,570.03	370.09
pccons2	34,063	39,015.90	488.31
pcconsWEEK2	34,063	24,494.23	306.56
Following the Deaton and Zaidi (2002) guidelines			
pcconsDZ	34,063	46,350.95	580.11
pcconsWEEKDZ	34,063	31,829.28	398.36
Following the CBS (2011) guidelines			
pcconsNLSS	34,141	37,879.21	474.08
Following the Furuta (2016) guidelines			
pcconsFURUTA	34,141	44,022.19	550.97
WDI (2011) calculations for Nepal			
Per capita consumption expenditure (constant 2010 USD)***		36,965.94	462.65
CBS (2011) calculations using NLSS III			
Per capita consumption expenditure, current NRs. (USD), all obs.		34,829.00	436.44

Variable name	Obs.	mean	mean (USD*)
Author's "optimised" income aggregation methodology**			
pcincome	**34,146**	**97,028.35**	**1,214.37**
pcincome1	34,146	97,118.68	1,215.50
pcincome2	34,146	99,669.48	1,247.43
pcincome3	34,146	48,614.26	608.44
pcincome4	33,590	89,692.54	1,122.56
WDI (2011) calculations for Nepal 2011			
Adjusted net national income per capita (constant 2010 USD)***		51,359.72	642.80
CBS (2011) calculations using NLSS III			
Per Capita Income, current NRs. (USD), all obs.		41,659.00	521.39

Note: *totcons* is total household consumption and *pcpcons* is per-capita annual consumption. Both utilise the "typical month" criteria. *totconsWEEK* is the total household consumption utilising the "past week" criteria, while *pcpconsWEEK* is the per-capita annual consumption using the "last week" criteria".

Note: *totincome* is total household income and *pcincome* is per-capita annual income.

* Average exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

** Consumption aggregation methodology combining the recommendations from Deaton and Zaidi (2002), CBS (2011) and Furuta (2016)

*** WDI (2011): First, household consumption is computed, then the per capita values are calculated. Household final consumption expenditure is the market value of all goods and services, including durable products purchased by households. It excludes purchases of dwellings, but includes imputed rent for owner-occupied dwellings. It also includes payments and fees to governments to obtain permits and licenses. Here, household consumption expenditure includes the expenditures of non-profit institutions serving households, even when reported separately by the country. Data sources: World Bank national accounts data, and OECD National Accounts data files.

* Average exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

**Largely based on CBS (2011, Ch. 11) but also adjusted to comments by Coulombe and McKay (2008), Food and Agriculture Organisation (2008), Quinones et al. (2009), International Labour Organisation (2003) and Furuta (2016).

***WDI (2011), Adjusted net national income is GNI minus consumption of fixed capital and natural resources depletion.

2.6.2. Comparison of the Distribution and Disparity of Income and Consumption Expenditure

This section describes and briefly tries to explain the distribution and disparity for both income and consumption using primarily the main reference aggregates as well as some of the other aggregates. This dissertation focuses on four main ways to do so.

1. **Lorenz Curves** (Figure 2.5): graphical illustrations
2. **Percentile Ratios** (Table 2.15)
3. **Income and consumption distributions by population** (quintiles; Table 2.16 and Table 2.17)
4. **Income and consumption distributions by geographical zones** (Urban vs. rural, ecological zones, Development Regions; Table 2.18 through to Table 2.23)

Complementary to the description of the distributional inequality above, the next Section (2.6.3) details the computation of the statistical inequality in outcome indices.

1. Lorenz curves (Figure 2.5)

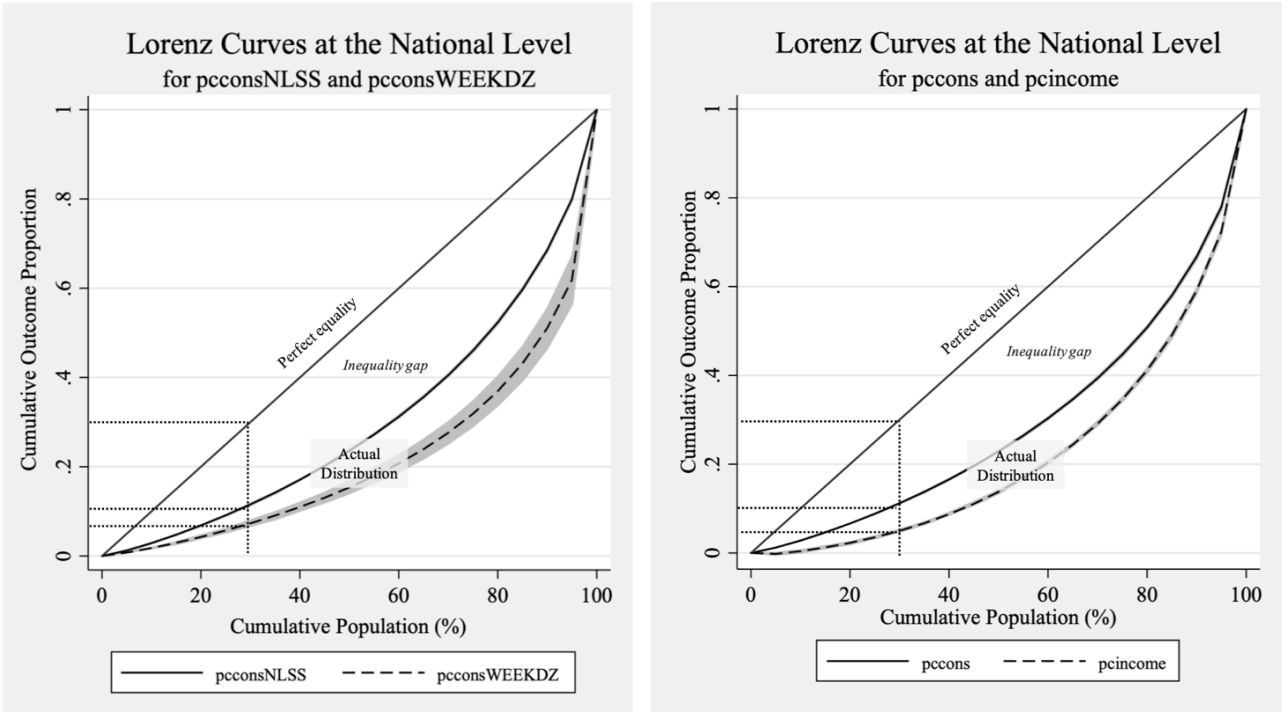
Here, the graphical illustrations of two selected Lorenz curve graphs serve as a first point of apercu to look at distributional inequality in Nepal at the national level. The left graph represents two Lorenz curves, one for per capita consumption using the CBS' (2011) methodology and the other using Deaton and Zaidi's (2002) weekly criterion methodology. The right graph also illustrates two Lorenz curves using the author's suggested reference variables for income and consumption. From both graphs it can be seen that a) the use of either income or consumption, and b) the use of different outcome aggregates following different aggregation methodologies, has different effects on distribution and disparity.

The **left graph** represents two Lorenz curves. One for per capita consumption following the CBS' aggregation methodology, and one for per capita consumption using Deaton and Zaidi's (2002) weekly criterion aggregation methodology. These two consumption aggregates were chosen for this graph because they were the furthest apart from each other amongst all the consumption aggregates. This is due to a few reasons. A) The CBS methodology cuts out the bottom and top 1%, meaning that a significant number of the top consuming observations (who have a disproportionately larger share of national consumption), and the poorest of the poor (who have a disproportionately smaller share of national consumption) were eliminated. This makes the overall consumption distribution appear statistically less unequal. That is, talking in terms of the Lorenz curve, the inequality gap to perfect equality is reduced. The Lorenz curve for *pcconsNLSS* is closer to the perfect equality line, illustrating that there is less consumption disparity across the population.

B) *PcconsWEEKDZ* uses the weekly criterion, wherever possible, in the aggregation process for consumption. As previously stated, this tends to underestimate the consumption estimate (compared to *pcconsDZ,* monthly criterion in Table 2.14), because less frequently and seasonally consumed items, such as rice bulk buying or education expenses, are not recorded. This disproportionately increases consumption for the middle and upper income households, as can be seen by the Lorenz curve (i.e. further away from the line of perfect equality and an increasingly steeper slope as one passes the 50% cumulative population share). For instance, as previously explained, Nepal has a dual education system where wealthier families try to send their children to private school. From interviews with people on the ground, the author realised that tuition fees as either paid at the monthly or the bi-weekly basis. Also, these wealthier parents tend to spend more money for their children's cafeteria school lunches, etc. These factors are more likely to be captured, more disproportionately for the wealthier families, using the weekly criterion. Public schools are mainly free, and often children bring their own lunches to school, or classes finish before eating hours, and other school-related expenses (school uniforms) are only lump sum expenses. These factors are amongst those that downward-bias schooling expenditure using the weekly criterion disproportionately for the individuals in the lower segment of the distribution, i.e. worsening consumption disparity. Furthermore, poorer families are dependent on bulk-buying the most basic foods, such as rice, and often they are bought (or received in-kind) between large time intervals. Also, many families home-grow their rice and therefore only a yearly value for this is computed. Richer families, however, who tend to work less in agriculture and receive less in-kind payments, can afford to buy rice at the markets at regular intervals, upward biasing the weekly criterion consumption aggregate. Again, this worsens the statistical image of the consumption disparity in the country. So, the left graph shows that a) the distributional disparity between the consumption aggregates depends on the aggregation methodology followed, and b) that the weekly criterion tends to disproportionately downward-bias consumption for the lower income groups and upward bias for the upper income groups.

The **right graph** illustrates the two Lorenz curves for the author's suggested reference variables for income (*pcincome*) and consumption (*pccons*). Section 2.1 elaborated on the literature on the measurement of welfare, and particularly income and consumption expenditure as proxies for economic welfare. As expected, distributional inequality is larger for income than for consumption. The income Lorenz curve is further away from the line of perfect equality than the one for consumption. The inequality gap is larger, the slope for the lower segments of the population is less steep than for consumption, and the slope for the upper part of the distribution is much steeper than that for consumption. Judging roughly from the graph, one can already induce that the bottom 30% of the population earns less than 5% of the country's overall income, while for overall consumption this is about 10%. That is, the share of overall economic outcome of the poorest 30% of the country for consumption is more than double that for income.

Figure 2.5: Lorenz Curves



Source: Author's computations.

2. **Percentile ratios** (Table 2.15)

Table 2.15 shows some selected percentile ratios for the various computed income and consumption aggregates. **Percentile ratios** are the relative distance between two points considered, depending in the top, the lower or the middle range of the income distribution. Upfront and as expected, the percentile ratios for income are always higher than for consumption. The only exception being the p10/p50 percentile ratio. This is because it measures the ratio of the bottom 10% of the distribution compared to the median i.e. looking exclusively at the lower parts of the distribution (bottom 50%). For this, the consumption value is higher than for income because the distribution of consumption gradually increases and is more equal (also in the lower parts of the distribution) than for income (see the Lorenz curve comparison in Figure 2.5). So, in terms of consumption, an increase in economic welfare can be noted within the lower 50% of the distribution towards the distribution median. For income, however, there is not much difference between the bottom 10% earners and the earners at the median level. This is because at the income level, a) the society is a lot more unequal, where most income is earned by the top few percent earners, and b) a relatively small proportion of the country

has regular incomes, Nepal being an underdeveloped country. This means that when only looking at distributional differences of income of the bottom 50% of the population, they are all more or less "equally" bad off (see almost horizontal slope for the lower 50% cumulative population share of the Lorenz curve in Figure 2.5).

For the remaining reported percentile ratios, the income ratios are higher than those for consumption. This also goes for p90/p50 (i.e. focus on the upper 50% part of the distribution), which compares the top 10% earners (or consumers) to the median. Following the argument from p10/p50, the top 10% hold a much larger share of economic outcome than the rest of the country, so that when they are compared to the median, the inequality level of the top 10% compared to the top 50% is very high. And that is all the more important for income earner than for consumers.

The p90/p10 in the table represents the comparison of the 90^{th} to the 10^{th} percentile of the outcome distribution, i.e. how much the top 10% of the population hold compared to the bottom 10%. With a factor of 14.7 for the reference income aggregate, the level of unequal distribution is the highest amongst all given percentile ratios (as expected). It is also about 3 times that of the reference consumption variable.

P75/p25 compares the distributional inequality level of the middle class, suggesting that within group distribution inequality is less when using consumption over income as the outcome variable. This again suggests that income distribution is more skewed towards the top end of Nepalese society, while consumption seems to follow more of a normal, centralised distribution pattern.

Furthermore, a stark difference across all percentile ratios is noticed depending on which aggregates are used. For instance, for the consumption aggregate and the p90/p10 ratio, distribution inequality levels vary from 5.2 when using per capital consumption excluding education and health expenses (monthly criterion) compared to 8.1 when using Deaton and Zaidi's (2002) weekly criterion aggregation methodology. This suggests again that the aggregation methodology used to compute the economic outcome variables is crucial and directly impacts on its distributional inequality.

Table 2.15: Percentile Ratios using the Various Income and Consumption Expenditure Aggregates

	Percentile Ratios			
	p90/p10	p90/p50	p10/p50	p75/p25
Per capita income				
Per capita disposable income (pcincome)	15.334	3.898	0.254	4.432
	(0.196)	(0.038)	(0.002)	(0.039)
Per capita income (pcincome1)	15.350	3.901	0.254	4.432
	(0.182)	(0.040)	(0.003)	(0.038)
Per capita income including utilities (pcincome2)	14.421	3.824	0.265	4.293
	(0.174)	(0.030)	(0.003)	(0.035)
Per capita income excluding wage income (pcincome3)	10.648	3.978	0.374	3.461
	(0.153)	(0.048)	(0.003)	(0.031)
Per capita income with CBS (2011) outlier constraint (pcincome4)	13.974	3.822	0.274	4.302
	(0.246)	(0.042)	(0.004)	(0.037)
Per capita consumption expenditure				
Per capita consumption expenditure including education and health, typical month (pccons)	5.631	2.823	0.501	2.480
	(0.059)	(0.025)	(0.003)	(0.015)
Per capita consumption expenditure including education and health, weekly (pcconsWEEK)	7.845	3.640	0.464	2.939
	(0.068)	(0.031)	(0.003)	(0.022)
Per capita consumption expenditure excluding education and health, typical month (pccons2)	5.228	2.724	0.521	2.400
	(0.047)	(0.022)	(0.003)	(0.013)
Per capita consumption expenditure excluding education and health, weekly (pcconsWEEK2)	7.217	3.552	0.492	2.765
	(0.083)	(0.040)	(0.003)	(0.019)
Per capita consumption expenditure, typical month, Deaton and Zaidi (2002), (pcconsDZ)	5.630	2.845	0.505	2.492
	(0.053)	(0.026)	(0.003)	(0.019)
Per capita consumption expenditure, weekly, Deaton and Zaidi (2002), (pcconsWEEKDZ)	8.160	3.730	0.457	3.047
	(0.074)	(0.031)	(0.003)	(0.033)
Per capita consumption expenditure, CBS (2011), (pcconsNLSS)	5.533	2.823	0.510	2.503
	(0.045)	(0.023)	(0.003)	(0.013)
Per capita consumption expenditure, Furuta (2016), (pcconsFURUTA)	5.745	2.928	0.510	2.515
	(0.060)	(0.029)	(0.003)	(0.016)

Source: Author's calculations. *Note*: Bootstrapped standard errors with 100 reps. in brackets.

3. **Income and consumption distributions by population** (quintiles; Table 2.16 and Table 2.17)

Table 2.16 shows the descriptive statistics for per capita income distribution across the population in quintiles using the various different income aggregates. Table 2.17 shows the same for per capita consumption expenditure. The tables report various complementary descriptive statistics, such as the standard deviations, median, interquartile range, and the Coefficient of Variation.

First, in general, the distribution division by quintiles shows a higher degree of inequality for the income than for the consumption aggregates. This goes in line with the descriptions made using the Lorenz curves and the percentile ratios. For instance, the bottom 20% of the population earns less than 3% of overall income compared to 56% of the top 20% earners when using the author's proposed reference income variable. For consumption, this is at 9% against 42%. This goes in hand with the literature on the measurement of economic welfare when using income or consumption as proxies.

Second, it is important to highlight that depending on which consumption (or income) aggregate is used, the distribution of outcome changes. For instance, using *pccons*, the bottom 20% of the country consume 9%, while the top 20% consume over 42% of overall consumption. When comparing it to *pcconsDZ*, this is over 5% versus 58%, respectively. So, the choice of aggregation methodology and which aggregate(s) is used in the econometric analysis is important and can thus have a significant impact on the results, particularly when conducting an inequality analysis which is based on the

distributional characteristics of economic resources. It also highlights that knowing how the variables were aggregated is crucial when interpreting the results.

Third, referring to the bottom quintile, the income aggregates for *pcincome, pcincome1*, *pcincoem2*, and *pcincome3* show large negative minimum values, relatively high standard deviation, and a high estimate of the Coefficient of Variation. CBS' (2011) *pcincome4* aggregate displays more "normal" values for this because the CBS mechanically eliminated all observations in the bottom 1% of the distribution, considering all of them as outliers. However, in this dissertation, the outliers were studied individually at every level (component) of the aggregation process. It was the sum of the negative values (which were individually judged as acceptable) that rendered some households (and individuals) to have such high negative income levels. Particularly, the farm income components bore large negative income values. Within the overall farm income, especially crop income (agricultural expenses outweighing crop income, e.g. through seed and fertilizers), net livestock and net land rental income induced the large negative values. Furthermore, non-farm enterprise income and transfer income has some households with large negative observations. It appears that some of the poorest households further indebt themselves to allow family members to emigrate in order to help lift them out of extreme poverty. These are some of the key components that are at the origin of the large Coefficient of Variation and standard deviations for the bottom income quintile. Consumption aggregates can never be negative, just extremely low. So the bottom quintile for *pccons* shows that there are some individuals whose annual consumption levels are less than 5,700NRs (= ca. 71USD).

Fourth, the disparity within the top quintiles for both consumption and income is very high (as already shown using the percentile ratios). For instance, the maximum value of the top consumers for *pccons* is 742,000NRs (= ca. 9,300USD), which is almost 10 times the median value of the group. Also the Coefficient of Variation is high at over 0.62. But for *pcincome*, this is even more extreme. The maximum value for top earners is at 4,437,000NRs (= ca. 55,500USD), which is almost 20 times the median value of the group. The quintile has a Coefficient of Variation of over 0.82.

The brief descriptive comparison of economic outcome by population quintiles clearly shows a) the difference between the utilisation of income or consumption as an economic welfare proxy, b) the importance of carefully choosing the most adequate aggregation methodology for the outcome variable, and c) the great distributional inequality in Nepal in economic welfare.

Table 2.16: Nominal Per Capita Income by Population Quintile
for the Various Income Aggregates

Quintiles	Mean per capita income (NRs.)	Share of total income (%)	Cumulative share of income (%)	Obs.	Min.	Max.	Standard Deviation	Median	Interquartile Range	Coefficient of Variation
pcincome										
Poorest (First)	11,789.21	2.60	2.60	7,314.00	-2,796,227.00	23,276.00	60,952.76	14,581.83	8,411.28	5.17
Second	33,237.13	7.13	9.74	7,112.00	23,295.86	44,789.82	6,324.74	32,548.88	11,134.80	0.19
Third	60,415.40	12.75	22.49	6,994.00	44,800.91	78,732.14	9,724.28	59,811.96	16,683.81	0.16
Fourth	108,851.70	21.69	44.18	6,601.00	78,771.71	150,609.00	20,170.43	105,712.90	34,147.45	0.19
Richest (Fifth)	301,950.30	55.82	100.00	6,125.00	150,662.30	4,437,499.00	251,901.90	231,339.00	135,275.50	0.83
Nepal mean	**97,028.35**	**100.00**		**34,146**	**-2,796,227.00**	**4,437,499.00**	**150,070.40**	**56,114.50**	**90,671.00**	**1.55**
pcincome1										
Poorest (First)	11,807.69	2.60	2.60	7,314.00	-2,796,227.00	23,284.00	60,905.15	14,588.95	8,410.41	5.16
Second	33,248.47	7.13	9.73	7,112.00	23,295.86	44,793.40	6,325.18	32,551.75	11,122.60	0.19
Third	60,433.87	12.74	22.47	6,991.00	44,800.91	78,856.75	9,716.70	59,878.20	16,706.47	0.16
Fourth	108,940.40	21.70	44.18	6,607.00	78,883.00	150,856.40	20,202.02	105,726.80	34,366.00	0.19
Richest (Fifth)	302,372.90	55.82	100.00	6,122.00	151,012.30	4,437,537.00	252,062.10	231,536.80	136,054.50	0.83
Nepal mean	**97,119.68**	**100.00**		**34,146**	**-2,796,227.00**	**4,437,537.00**	**150,185.00**	**56,147.00**	**90,784.41**	**1.55**
pcincome2										
Poorest (First)	12,111.96	2.62	2.62	7,350	-2,769,797.00	24,081.00	61,844.49	15,617.71	8,790.86	5.11
Second	34,546.63	7.27	9.88	7,161	24,082.00	46,786.83	6,563.64	34,495.00	11,539.80	0.19
Third	62,580.54	12.78	22.67	6,952	46,821.33	82,123.75	10,239.82	61,416.75	18,191.00	0.16
Fourth	112,735.90	22.01	44.68	6,644	82,169.75	154,962.50	20,925.25	110,101.30	36,395.13	0.19
Richest (Fifth)	311,777.70	55.32	100.00	6,039	155,046.20	4,439,589.00	259,750.10	239,449.30	149,592.30	0.83
Nepal mean	**99,669.48**	**100.00**		**34,146**	**-2,769,797.00**	**4,439,589.00**	**153,882.50**	**57,209.17**	**91,329.67**	**1.54**
pcincome3										
Poorest (First)	6,663.20	3.31	3.31	8,236	-2,796,227.00	14,647.57	59,976.74	10,329.39	4,800.20	9.00
Second	18,847.20	8.52	11.83	7,507	14,650.83	23,842.67	2,612.56	18,630.24	4,473.00	0.14
Third	30,524.64	12.79	24.62	6,957	23,863.57	39,065.67	4,360.64	29,950.00	7,373.40	0.14
Fourth	53,887.77	19.80	44.42	6,100	39,085.00	75,457.34	10,517.12	52,409.33	16,940.92	0.20
Richest (Fifth)	172,567.00	55.58	100.00	5,346	75,519.74	2,517,734.00	142,701.50	123,962.40	97,058.83	0.83
Nepal mean	**48,614.26**	**100.00**		**34,146**	**-2,796,227.00**	**2,517,734.00**	**84,716.56**	**26,003.75**	**36,311.80**	**1.74**
pcincome4										
Poorest (First)	15,538.37	3.79	3.79	7,340	3,551.43	23,952.17	5,122.73	15,792.00	8,575.00	0.33
Second	34,250.68	8.05	11.83	7,078	23,982.67	46,216.00	6,489.56	33,833.00	10,986.83	0.19
Third	61,309.96	13.80	25.63	6,781	46,256.17	80,428.80	9,561.20	60,390.80	16,262.50	0.16
Fourth	109,110.70	23.37	49.01	6,454	80,466.00	149,710.30	19,795.72	106,229.50	34,843.85	0.18
Richest (Fifth)	258,775.70	50.99	100.00	5,937	149,713.70	659,041.00	106,954.30	223,200.40	125,311.90	0.41
Nepal mean	**89,692.54**	**100.00**		**33,590**	**3,551.43**	**659,041.00**	**96,268.47**	**55,714.25**	**88,617.83**	**1.07**

Source: Author's calculations.

Table 2.17: Nominal Per Capita Consumption Expenditure by Population Quintile for the Various Consumption Expenditure Aggregates

Quintiles	Mean per capita consumption (NRs.)	Share of total consumption (%)	Cumulative share of consumption (%)	Obs.	Min.	Max.	Standard Deviation	Median	Interquartile Range	Coefficient of Variation
pccons										
Poorest (First)	15,705	9.20	9.20	8,795	5,689.14	20,462.12	3,224.82	16,289.76	4,928.18	0.21
Second	24,638	12.79	21.99	7,796	20,463.21	29,495.14	2,658.19	24,440.55	4,511.92	0.11
Third	35,304	15.71	37.69	6,683	29,496.15	42,477.82	3,714.67	35,060.24	6,297.42	0.11
Fourth	53,869	20.68	58.38	5,766	42,485.48	70,132.18	7,841.11	52,410.82	12,834.26	0.15
Richest (Fifth)	124,457	41.62	100.00	5,023	70,198.51	741,725.60	76,811.45	98,560.20	52,262.95	0.62
Nepal mean	**44,092**	**100.00**		**34,063**	**5,689.14**	**741,725.60**	**46,537.52**	**30,111.22**	**29,891.09**	**1.06**
pcconsWEEK										
Poorest (First)	8,273	7.17	7.17	8,729	2,777.31	10,895.55	1,703.59	8,430.67	2,659.74	0.21
Second	13,471	10.37	17.54	7,756	10,897.20	16,452.44	1,581.27	13,323.25	2,777.57	0.12
Third	20,349	13.38	30.93	6,625	16,455.23	25,653.21	2,619.13	19,878.41	4,442.51	0.13
Fourth	34,226	19.62	50.55	5,774	25,687.71	46,806.63	6,098.40	33,187.17	10,201.26	0.18
Richest (Fifth)	96,179	49.45	100.00	5,179	46,810.44	721,115.10	71,750.89	71,702.98	46,066.27	0.75
Nepal mean	**29,570**	**100.00**		**34,063**	**2,777.31**	**721,115.10**	**40,762.86**	**17,018.93**	**20,903.36**	**1.38**
pcconsDZ										
Poorest (First)	14,884	5.06	5.06	1,198.00	4,497.89	19,187.26	3,047.02	15,505.10	4,427.07	0.20
Second	23,470	7.98	13.04	1,197.00	19,190.67	27,947.13	2,593.53	23,379.46	4,538.63	0.11
Third	33,687	11.46	24.50	1,198.00	27,949.68	40,559.91	3,621.23	33,362.71	6,251.61	0.11
Fourth	51,575	17.53	42.02	1,197.00	40,597.82	66,851.74	7,455.94	50,572.94	12,143.52	0.14
Richest (Fifth)	170,591	57.98	100.00	1,197.00	66,859.60	3.00E+07	1,013,445.00	97,024.02	55,748.94	5.94
Nepal mean	**58,830**	**100.00**		**5,987**	**4,497.89**	**30,000,000.00**	**456,613.90**	**33,358.43**	**36,153.48**	**7.76**
pcconsNLSS										
Poorest (First)	13,952	9.48	9.48	8,785.00	4,280.58	18,046.61	2,879.74	14,458.48	4,298.49	0.21
Second	21,744	13.16	22.64	7,828.00	18,051.51	25,983.11	2,334.02	21,594.98	4,062.52	0.11
Third	31,100	16.00	38.64	6,655.00	26,011.88	37,579.43	3,348.61	30,797.88	5,881.65	0.11
Fourth	47,682	21.71	60.35	5,887.00	37,579.92	61,372.29	6,785.65	46,649.48	11,303.40	0.14
Richest (Fifth)	102,844	39.65	100.00	4,986.00	61,387.52	548,502.20	49,941.62	86,562.98	43,040.38	0.49
Nepal mean	**37,879**	**100.00**		**34,141**	**4,280.58**	**548,502.20**	**35,023.22**	**26,553.68**	**26,815.59**	**0.92**
pcconsFURUTA										
Poorest (First)	13,941.91	8.15	8.15	8,781.00	4,276.58	17,977.12	2,873.20	14,445.22	4,298.36	0.21
Second	21,740.79	11.33	19.47	7,830.00	17,982.95	25,969.11	2,342.38	21,598.86	4,113.23	0.11
Third	31,122.60	13.76	33.23	6,646.00	25,976.01	37,676.00	3,365.18	30,826.95	5,844.15	0.11
Fourth	47,845.24	18.59	51.83	5,841.00	37,676.62	61,899.03	6,888.81	46,816.46	11,409.64	0.14
Richest (Fifth)	143,565.80	48.17	100.00	5,043.00	62,083.95	3.00E+07	679,616.50	88,375.71	49,090.07	4.73
Nepal mean	**44,022**	**100.00**		**34,141**	**4,276.58**	**30,000,000.00**	**264,711.00**	**26,549.68**	**26,998.07**	**6.01**

Source: Author's calculations. *Note*: not all per capita consumption expenditure variations are given.

4. **Income and consumption distributions by geographical zones** (urban vs. rural, ecological zones, Development Regions; Table 2.18 through to Table 2.23)

Table 2.18 shows the descriptive statistics for per capita income distribution across the country by geographical zone using the various income aggregates. Table 2.19 shows the same for per capita consumption expenditure. The purpose of these two tables is to first highlight, again, the importance for choosing the adequate aggregation methodology, because the methodology process will affect the distribution of outcome also at the geographical level. The second purpose is to illustrate the divergences between the income and consumption aggregates. The third purpose is to show the great differences in economic disparity across regional regions.

First, looking at the **consumption aggregates** and the Development Region stratification, for *pccons*, the ratio of national mean consumption in the Central region is 130% and in the mid-western region is 62%. However, for *pcconsWEEKDZ*, these are 153% and 47%, respectively. Given that the Central region is the economic centre of Nepal and hosts the capital and some SEZs, the ratio of mean consumption compared to the overall national mean consumption level is expected to be higher. Also, the mid-western region, being predominantly rural, agricultural, and suffering from youth outmigration, is expected to have much lower mean consumption levels compared to the national mean.

Second, interestingly, the **consumption distributions between the rural and urban areas are slightly more unequal than the income** ones using this dissertation's reference variables. For instance, for *pccons*, the ratio of the mean consumption in urban areas is 167% of the national mean consumption level, while in rural areas it is only 87%. This disparity is more moderate for *pcincome* with 158% and 89%, respectively, yet they are far apart. One possible explanation lies in the breakdown of wage income. While wage income makes up over 50% of overall income (51% in rural and 47% in urban areas), Table 2.22 indicates that a much larger proportion of the rural wage income comes from daily labour (30%) compared to urban areas (5%). As one would expect, the share of non-agricultural long-term wages in urban areas is much higher than in rural areas (85% versus 62%). So, it is a matter of the component share of income and its sub-component share for wage income being higher in rural areas and therefore representing a higher income share in rural and provincial regions compared to urban centres and the more developed Central region. The distribution disparity between income and consumption are more similar for the **Ecological and Development Region zones**. This is primarily because their boundaries serve a public administration purpose. Thus, these areas include both urban and rural centres, as well as regions of different development statuses, hence blurring the regional disparity for the Development Region and Ecological Zone categories.

Third, referring to *pccons* and the **ecological zone stratification**, the Terai region has the highest mean ratio compared to the national mean, followed by the mountain and hill region, which are much more rural and underdeveloped (114%, 93%, and 81%, respectively). The **urban rural cleavage** in distributional disparity is the highest, with urban mean consumption being over 168% of the national average, while rural areas see a mean consumption level of barely 87% of the national mean. Possibly, the urban rural divide is the most reflective of the distributional outcome inequalities in Nepal, suggesting that higher welfare is to be expected in the urban centres. This strong pull factor (Lee, 1966) is most certainly an essential element when trying to explain Nepal's high domestic migration flows. The chapter on the results interpretation of Inequality of Opportunity, however, will highlight A) that this urban welfare aspiration can be misleading, particularly for the most opportunity deprived populations who face even higher discrimination in urban areas, and B) that while consumption and income levels are higher in the urban areas, living expenses are also higher so they do not necessarily drive better living standards.

Table 2.20 shows the component shares for the total per capita **income aggregate**, and Table 2.22 complements this with a componential share breakdown for wage income. In both tables, *pcincome* serves as the reference variable, and the distributions are disaggregated by geographical zone.

At the national level, more than 50% of per capita income is wage income. Of this, more than 67% is from long-term non-agricultural wages and 23% from daily wages. A further 9% is attributed to contract and piece rate wages and 1% to agricultural long-term wages. There are some stark geographical variations in the wage compositions. For instance, as one would expect, daily wage income in rural areas is higher than in urban areas, at 30% and 5%, respectively. This goes in line with the Central region, which hosts the capital and other urban centres, to have only 20% in daily wages, compared to, for instance, the mid-western region (predominantly rural and agricultural) with over 35%. This is also in line when looking at the ecological zones. In the Terai region, which hosts both the capital and a number of Special Economic Zones, daily wages make up only 21% compared to 27% and 29% in the hill and mountain regions, respectively. Long-term non-agricultural wages make up 85% of the nation-wide urban and only 62% of the nation-wide rural wages.

Despite wage income making up to 50% of overall income, Table 2.20 shows that over 20% of total per capita income is still from farm income. Non-farm enterprise income makes up about 11%, transfer income about 9%, and housing income about 7% of the overall nation-wide per capita income. Again, there are some vast differences between geographical regions. The hill ecological zone, for instance, is the most dependent on transfer income, with a share of over 10%. For the Development Region stratification level, the western region depends on transfer income for over 16%. Furthermore, over 17% of urban income and 13% of income in the Central region arises from non-farm enterprise income. This is very high compared to the rural areas and the other Development Regions. It suggests that these are the areas with the highest variety of economic activity.

Table 2.21 shows the component shares for the total per capita **consumption aggregate** and Table 2.23 complements this with a componential share breakdown for food consumption. In both tables, *pccons* serves as the reference variable, and the distributions are disaggregated by geographical zones. At the national level, **food** consumption is by far the largest consumption component at over 41%. However, the more remote and rural an area becomes, the larger the share of the food component of the overall consumption aggregate. In urban areas, food makes up only 30% of overall consumption, versus 47% in rural areas. At the Development Region level, in the central region the food component takes the smallest share (37%), while for the mid- and far-western regions it is over 53% and 49%, respectively. This is in line with Engel's Law (see Appendix 2:). These are Nepal's most impoverished regions and HHs are by enlarge dependent on the home production of food. Table 2.23 shows that out of overall food consumption, the mid- and the far-western regions home produce almost 50% of their food consumption, while this is as low as 22% in the central region, and only 12% in urban areas. This is in line with the descriptive statistics showing that in urban centres, more than 84% of food consumption is purchases due to A) the readily availability of food produce and markets, and B) the monetary purchasing power of urban HHs. Furthermore, it is in the most rural areas, such as the far-west region or the mountain areas, where the in-kind received food component is the largest at 6%.

The **housing share** of consumption is highest in urban compared to rural areas (24% and 13%, respectively), where real-estate prices have significantly risen over the past years. This is particularly due to the inflow of remittances allowing local families to build their own houses.

As discussed in Section 2.4.1, the inclusion of utilities is problematic due to variation in the geographical coverage of access to utilities, in quality, and in price subsidisation. Table 2.21 reflects this. While utility consumption makes up over 6% of urban consumption, in the remote region of the far-west for instance, it is only 2%.

Education is another important consumption component with significant geographical disparity. Urban areas see much higher education consumption than rural areas (9% and 6%, respectively). This is A) due to the availability of private schools and higher education institutions in urban centres, which require more financial commitment (public education is free of charge until for the first 8 years of education, equivalent to middle school). B) More financially affluent families live in urban centres. While they spend only 9% of their disposable income on education, this 9% is much

larger in absolute terms than for poor rural families. For poor rural families, any additional educational expense implies immediate consumption and welfare trade-offs. Thus, Section 2.4.1 elaborated on the literature debate on whether educational expenses should be included or excluded from the overall consumption aggregate.

Looking at the **durable goods** components (purchased over the past year or older), there is some significant divergence between urban and rural areas. Urban HHs allocate more that 6-8% of their total consumption to durable goods, while rural HHs only about 3-4%. In the mid-western region, the consumption of durables is particularly low (2%). This can be due to several reasons. For instance, the greater availability of durable goods and the variety of them in urban centres, or the higher purchasing power of urban HHs allowing them to consume a larger proportion of their overall consumption of durable goods.

Overall, the descriptive analysis of the distributional disparity in economic outcome by geographical region showed, first, that the choice of aggregation methodology for the economic outcome variable impacts on the distributional disparity by geographic region. Second, it showed that the levels of economic outcome disparity vary and are high in amplitude. Third, it showed that the urban rural comparison shows the most extreme divides compared to the other geographical groupings. Fourth, the geographical decomposition complements the demographic decomposition when analysing the distributional disparity of economic outcome.

This section described and briefly explained the distribution and disparity of economic outcome by using graphical illustrations (the Lorenz curve), and descriptive statistics. The comparison of the percentile ratios and the analysis of the distribution by population quintiles, as well as geographical zones, complemented the disparity description of the graphical illustrations. The next Section (2.6.3) details the computation of the statistical inequality in outcome indices.

Table 2.18: Per Capita Income Distribution by Geographical Zones Using the Different Income Aggregates

	pcincome			pcincome1			pcincome2		
	Per capita mean income (NRs.)	**Ratio of national mean income (%)**	**Obs.**	**Per capita mean income (NRs.)**	**Ratio of national mean income (%)**	**Obs.**	**Per capita mean income (NRs.)**	**Ratio of national mean income (%)**	**Obs.**
Ecological Zone Total	97,064.94	100.00	32,802	97,154.01	100.00	32,802	99,708.13	100.00	32,786
Mountains	98,306.80	101.28	3,083	98,352.36	101.23	3,083	100,912.44	101.21	3,093
Terai	107,308.53	110.55	17,784	107,431.17	110.58	17,784	111,452.96	111.78	17,740
Hill	81,480.47	83.94	11,935	81,530.76	83.92	11,935	81,965.44	82.21	11,953
Urban/Rural Total	97,064.94	100.00	32,802	97,154.01	100.00	32,802	99,708.13	100.00	32,786
Urban	153,630.77	158.28	5,234	153,855.08	158.36	5,234	157,260.76	157.72	5,221
Rural	86,325.47	88.94	27,568	86,388.87	88.92	27,568	88,807.26	89.07	27,565
Development Region Total	97,064.94	100.00	32,802	97,154.01	100.00	32,802	99,708.13	100.00	32,786
Eastern	87,808.55	90.46	7,503	87,863.86	90.44	7,503	89,125.65	89.39	7,471
Central	122,131.23	125.82	11,014	122,255.23	125.84	11,014	126,406.24	126.78	11,039
Western	98,554.18	101.53	6,642	98,640.19	101.53	6,642	101,335.03	101.63	6,634
Mid-Western	64,810.86	66.77	4,634	64,925.09	66.83	4,634	66,109.99	66.30	4,637
Far-Western	74,779.91	77.04	3,009	74,793.39	76.98	3,009	76,194.94	76.42	3,005
Nepal	**97,028**	**100**	**34,146**	**97,119**	**100**	**34,146**	**99,669.48**	**100.00**	**34,146**

	pcincome3			pcincome4		
	Per capita mean income (NRs.)	**Ratio of national mean income (%)**	**Obs.**	**Per capita mean income (NRs.)**	**Ratio of national mean income (%)**	**Obs.**
Ecological Zone Total	48,569.29	100.00	32,786	89,864.29	100.00	32,302
Mountains	44,565.91	91.76	3,093	84,310.27	93.82	3,032
Terai	53,777.85	110.72	17,740	100,785.19	112.15	17,392
Hill	41,874.94	86.22	11,953	75,291.43	83.78	11,878
Urban/Rural Total	48,569.29	100.00	32,786	89,864.29	100.00	32,302
Urban	79,897.71	164.50	5,221	139,261.89	154.97	5,094
Rural	42,635.47	87.78	27,565	80,615.86	89.71	27,208
Development Region Total	48,569.29	100.00	32,786	89,864.29	100.00	32,302
Eastern	45,042.86	92.74	7,471	82,928.69	92.28	7,388
Central	61,068.38	125.73	11,039	113,133.61	125.89	10,834
Western	54,472.40	112.15	6,634	92,973.21	103.46	6,534
Mid-Western	29,319.89	60.37	4,637	60,394.94	67.21	4,573
Far-Western	28,092.27	57.84	3,005	60,799.42	67.66	2,973
Nepal	**48,614.26**	**100.00**	**34,146**	**89,692.54**	**100.00**	**33,590**

Source: Author's calculations.

Table 2.19: Per Capita Consumption Expenditure Distribution by Geographical Zones Using the Various Consumption Aggregates

	pccons			pcconsWEEK			pccons2			pcconsWEEK2		
	Per capita mean consumption (NRs.)	**Ratio of national mean consumption (%)**	**Obs.**	**Per capita mean consumption (NRs.)**	**Ratio of national mean consumption (%)**	**Obs.**	**Per capita mean consumption (NRs.)**	**Ratio of national mean consumption (%)**	**Obs.**	**Per capita mean consumption (NRs.)**	**Ratio of national mean consumption (%)**	**Obs.**
Ecological Zone Total	44,118.16	100.00	32,730	29,603.71	100.00	32,730	39,045.69	100.00	32,730	24,531.25	100.00	32,730
Mountains	40,947.59	92.81	3,051	26,570.78	89.75	3,051	36,336.14	93.06	3,051	21,959.34	89.52	3,051
Terai	50,400.54	114.24	17,737	34,832.35	117.66	17,737	44,470.87	113.89	17,737	28,902.68	117.82	17,737
Hill	35,597.21	80.69	11,942	22,612.68	76.38	11,942	31,680.13	81.14	11,942	18,695.61	76.21	11,942
Urban/Rural Total	44,118.16	100.00	32,730	29,603.71	100.00	32,730	39,045.69	100.00	32,730	24,531.25	100.00	32,730
Urban	73,916.02	167.54	5,265	56,002.13	189.17	5,265	64,994.33	166.46	5,265	47,080.44	191.92	5,265
Rural	38,405.95	87.05	27,465	24,543.17	82.91	27,465	34,071.38	87.26	27,465	20,208.60	82.38	27,465
Development Region Total	44,118.16	100.00	32,730	29,603.71	100.00	32,730	39,045.69	100.00	32,730	24,531.25	100.00	32,730
Eastern	39,300.45	89.08	7,459	24,836.77	83.90	7,459	35,153.51	90.03	7,459	20,689.82	84.34	7,459
Central	57,505.95	130.35	11,028	40,349.40	136.30	11,028	50,520.86	129.39	11,028	33,364.31	136.01	11,028
Western	45,964.02	104.18	6,619	31,287.49	105.69	6,619	40,420.00	103.52	6,619	25,743.48	104.94	6,619
Mid-western	27,512.49	62.36	4,604	16,906.38	57.11	4,604	24,869.04	63.69	4,604	14,262.93	58.14	4,604
Far-western	28,399.43	64.37	3,020	17,804.63	60.14	3,020	25,355.77	64.94	3,020	14,760.97	60.17	3,020
Nepal	**44,091.71**	**100.00**	**34,063**	**29,570.03**	**100.00**	**34,063**	**39,015.90**	**100.00**	**34,063**	**24,494.23**	**100.00**	**34,063**
	pcconsDZ			**pcconsWEEKDZ**			**pcconsNLSS**			**pcconsFURUTA**		
Ecological Zone Total	46,570.30	100.00	32,700	32,060.24	100.00	32,700	37,926.41	100.00	32,784	44,281.82	100.00	32,794
Mountains	59,304.65	127.34	3,036	44,903.30	140.06	3,036	34,972.57	92.21	3,044	57,573.04	130.02	3,082
Terai	52,711.58	113.19	17,730	37,147.62	115.87	17,730	42,752.19	112.72	17,768	50,245.56	113.47	17,751
Hill	34,206.78	73.45	11,934	21,234.80	66.23	11,934	31,515.38	83.10	11,972	32,006.43	72.28	11,961
Urban/Rural Total	46,570.30	100.00	32,700	32,060.24	100.00	32,700	37,926.41	100.00	32,784	44,281.82	100.00	32,794
Urban	94,851.66	203.67	5,207	76,918.76	239.92	5,207	61,567.55	162.33	5,252	91,677.23	207.03	5,228
Rural	37,426.12	80.36	27,493	23,564.32	73.50	27,493	33,416.63	88.11	27,532	35,293.09	79.70	27,566
Development Region Total	46,570.30	100.00	32,700	32,060.24	100.00	32,700	37,926.41	100.00	32,784	44,281.82	100.00	32,794
Eastern	37,571.07	80.68	7,450	23,119.41	72.11	7,450	34,032.06	89.73	7,478	35,274.85	79.66	7,478
Central	66,192.81	142.14	11,032	49,022.05	152.91	11,032	49,371.25	130.18	11,021	63,370.66	143.11	11,031
Western	47,526.14	102.05	6,626	32,857.22	102.49	6,626	38,902.52	102.57	6,647	44,827.65	101.23	6,632
Mid-Western	25,727.48	55.24	4,584	15,125.06	47.18	4,584	24,363.99	64.24	4,644	25,259.04	57.04	4,634
Far-Western	26,550.02	57.01	3,008	16,048.43	50.06	3,008	24,394.09	64.32	2,994	24,843.84	56.10	3,019
Nepal	**157,922.00**	**100.00**	**34,063**	**31,829.28**	**100.00**	**34,063**	**37,879.21**	**100.00**	**34,141**	**44,022.19**	**100.00**	**34,141**

Source: Author's calculations.

Table 2.20: Component Shares of Per Capita Income by Geographical Region (in %)

	Per capita farm income	Per capita wage income	Per capita non-farm enterprises income	Per capita non-agricultural rental income	Per capita transfer income	Per capita value owner occupied housing	Per capita other income	Total
Ecological Zone (32,786 obs.)	20.01	50.08	10.78	0.55	8.68	7.43	2.46	100.00
Mountains (3,093 obs.)	18.68	54.98	10.36	0.78	6.85	7.21	1.14	100.00
Terai (17,740 obs.)	19.02	50.49	10.17	0.62	8.32	8.24	3.15	100.00
Hills (11953 obs.)	22.44	47.70	12.13	0.35	10.01	5.88	1.49	100.00
Urban/Rural (32,786 obs.)	20.01	50.08	10.78	0.55	8.68	7.43	2.46	100.00
Urban (5,221 obs.)	14.49	47.51	17.02	0.85	5.32	11.73	3.07	100.00
Rural (27,565 obs.)	21.84	50.93	8.71	0.45	9.80	6.01	2.26	100.00
Development Region (32,786 obs.)	20.01	50.08	10.78	0.55	8.68	7.43	2.46	100.00
Eastern (7,471 obs.)	23.47	48.36	11.14	0.52	9.34	5.75	1.42	100.00
Central (11,039 obs.)	18.87	50.28	12.96	0.67	5.94	9.07	2.21	100.00
Western (6,634 obs.)	18.12	44.81	8.68	0.58	15.51	7.40	4.91	100.00
Mid-West (4,637 obs.)	22.68	55.02	7.78	0.16	7.54	5.41	1.41	100.00
Far-West (3,005 obs.)	18.77	62.57	6.76	0.40	5.00	5.27	1.24	100.00
Nepal (34,146 obs.)	**20.02**	**50.01**	**10.96**	**0.55**	**8.65**	**7.38**	**2.43**	**100.00**

Source: Author's calculations. *Note*: For variable *pcincome*.

Table 2.21: Component Shares of Per Capita Consumption by Geographical Region (in %)

	Per capita food	Per capita tobacco	Per capita education	Per capita health	Per capita non-food	Per capita housing	Per capita utilities	Per capita durables	Per capita durables 1yr old	Per capita own production	Total
Ecological Zone (32,786 obs.)	42.77	0.95	6.76	4.73	11.86	16.15	4.81	5.42	3.79	2.75	100.00
Mountains (3,093 obs.)	46.93	1.12	5.77	5.49	11.06	14.74	3.93	3.54	2.35	5.08	100.00
Terai (17,740 obs.)	40.03	0.95	7.27	4.49	11.51	17.84	5.17	5.93	4.21	2.60	100.00
Hills (11953 obs.)	47.31	0.92	5.99	5.02	12.81	13.00	4.33	4.91	3.34	2.38	100.00
Urban/Rural (32,786 obs.)	42.77	0.95	6.76	4.73	11.86	16.15	4.81	5.42	3.79	2.75	100.00
Urban (5,221 obs.)	30.25	0.67	8.68	3.39	11.61	24.13	6.40	8.07	6.00	0.79	100.00
Rural (27,565 obs.)	47.39	1.06	6.06	5.23	11.95	13.20	4.23	4.45	2.98	3.47	100.00
Development Region (32,786 obs.)	42.77	0.95	6.76	4.73	11.86	16.15	4.81	5.42	3.79	2.75	100.00
Eastern (7,471 obs.)	48.72	0.82	5.56	4.99	12.03	12.85	4.08	4.48	3.15	3.31	100.00
Central (11,039 obs.)	37.61	1.01	7.74	4.40	11.88	19.34	5.73	5.97	4.49	1.82	100.00
Western (6,634 obs.)	41.55	0.83	6.72	5.35	11.40	15.92	5.05	6.52	4.01	2.66	100.00
Mid-West (4,637 obs.)	53.20	1.26	5.38	4.23	12.92	9.88	2.91	3.29	1.90	5.04	100.00
Far-West (3,005 obs.)	49.53	0.95	5.88	4.84	11.09	13.90	2.52	3.82	2.88	4.59	100.00
Nepal (34,146 obs.)	**40.67**	**0.88**	**6.54**	**4.73**	**11.58**	**18.41**	**5.32**	**5.53**	**3.97**	**2.37**	**100.00**

Source: Author's calculations. *Note*: For variable *pccons*.

Table 2.22: Wage Component Shares of Per Capita Income (in %)

	per capita daily wage income	Per capita long-term agriculture wage employment	Per capita long-term non-agriculture wage employment	Per capita piece-rate/contract income	Total
Ecological Zone (32,786 obs.)	23.82	0.11	67.04	9.03	100
Mountains (3,093 obs.)	29.31	-	65.56	5.12	100
Terai (17,740 obs.)	21.20	0.08	68.44	10.29	100
Hills (11953 obs.)	27.36	0.22	64.62	7.81	100
Urban/Rural (32,786 obs.)	23.82	0.11	67.04	9.03	100
Urban (5,221 obs.)	4.77	0.02	84.80	10.41	100
Rural (27,565 obs.)	29.71	0.14	61.54	8.61	100
Development Region (32,786 obs.)	23.82	0.11	67.04	9.03	100
Eastern (7,471 obs.)	27.47	0.28	66.79	5.46	100
Central (11,039 obs.)	19.99	0.02	69.99	10.01	100
Western (6,634 obs.)	25.05	0.15	71.98	2.82	100
Mid-West (4,637 obs.)	35.48	0.09	60.97	3.46	100
Far-West (3,005 obs.)	17.83	0.10	50.27	31.80	100
Nepal (34,146 obs.)	**23.58**	**0.11**	**67.15**	**9.16**	**100**

Source: Author's calculations. *Note*: For variable *pcincome*.

Table 2.23: Food Component Shares of Per Capita Consumption Expenditure (in %)

	Purchased	Home produced	In-kind	Total
Ecological Zone (32,735 obs.)	64.27	31.40	4.33	100.00
Mountains (3,051 obs.)	56.64	37.56	5.80	100.00
Terai (17,742 obs.)	65.32	30.79	3.89	100.00
Hills (11,942 obs.)	64.63	30.68	4.70	100.00
Urban/Rural (32,735 obs.)	64.27	31.40	4.33	100.00
Urban (5,265 obs.)	84.49	12.06	3.45	100.00
Rural (27,470 obs.)	59.51	35.95	4.54	100.00
Development Region (32,735 obs.)	64.27	31.40	4.33	100.00
Eastern (7,459 obs.)	61.41	35.11	3.48	100.00
Central (11,033 obs.)	73.64	21.63	4.73	100.00
Western (6,619 obs.)	63.28	32.94	3.79	100.00
Mid-West (4,604 obs.)	49.31	45.79	4.90	100.00
Far-West (3,020 obs.)	47.94	46.34	5.72	100.00
Nepal (34,086 obs.)	**64.63**	**31.08**	**4.29**	**100.00**

Source: Author's calculations. *Note*: For variable *pccons*.

2.6.3. Comparison of the Inequality in Outcome Measures for Income and Consumption Expenditure

This section proposes three ways to look at the results of the various inequality indices reported in Table 2.24.

1. Differences between the various inequality measures
2. Differences in results between the income and consumption inequality indices
3. Differences that the various aggregation methodologies have on the inequality indices

1. Differences between the various inequality measures

Table 2.24 illustrates the results for the various inequality in outcome indices. As expected, different indices give different results because each index places emphasis on different segments of the distribution. As a reminder, the higher the value of α in the Generalised Entropy $G(\alpha)$, the more

sensitive the index is on inequalities in the upper segments of the distribution (i.e. relatively wealthier individuals). The same goes for ε in the Atkinson indices A(ε), the more positive the value of ε, the more sensitive the measure to distributional inequalities in the upper segment. It is important to note that the Gini coefficient is slightly apart from the G(-1), G(0), G(1) and G(2). First, the G(-1) places the largest emphasis for distributional inequalities amongst the poorest of the poor and poor. As shown in the previous section, there are individuals in Nepal who are forced to survive on less than 80USD per year. Yet, the mean consumption for the bottom quintile is almost 200USD per capita. The standard deviation for the bottom quintile is also the largest for the lower 50% of the distribution for the entire population. This results in the relatively high G(-1) compared to the MLD.

Second, it is important to note here that the **MLD** estimate is lower than those of the other inequality indices. Despite placing a comparatively large emphasis on the lower parts of the distributions compared to the Gini coefficient, for instance, the estimate value of it is relatively low. This is because the emphasis lies on the overall bottom and the mid-bottom parts of the distribution, thus focussing less on the poorest of the poor like G(-1) and more on the lower middle distribution. Furthermore, Nepal is one of the poorest countries in the world, and a vast majority of its population lives in more or less equally low living conditions (i.e. see low slopes of the Lorenz curves above, or the marginally increasing mean estimates for economic outcomes for the first two to 3 bottom quintiles). Because the standard deviation of the lower parts of the distribution is so low (compared to the upper two quintiles), the MLD overall value is relatively low. Were the middle class (mid 25-75% of the population) to increase their economic welfare, then particularly the mid 25-50% changes in welfare would increase the MLD, given that the poorest of the poor remain in a stagnant situation. Thus, the MLD estimate is the lowest amongst the Generalised Entropy inequality indices. Yet, the MLD is judged to be the most suitable inequality in outcome index of the Generalised Entropy indices, to serve as the main inequality in outcome reference variable for the subsequent inequality of opportunity analyses. The theory behind this is described and explained in Section 2.2, and the statistical results here reconfirm the necessity of an inequality in outcome index, that is to serve as a reference for any inequality of opportunity analysis, to place emphasis on the lower parts of the distribution. Since the MLD values are somewhat lower than those of the other indices, it indicates that individuals in the lower-middle segment of the distribution are all more or less equally bad off (also, low variance and standard deviation in the lower segments of the population). So, any marginal difference in economic outcome has an important impact on the economic welfare of a household. The inequality of opportunity analysis tries to find out as to how much of this inequality in outcome is due to opportunity deprivation. So, an inequality in outcome index that is particularly sensitive to variations in the lower segment allows to more finely pinpoint a) differences in opportunity discrimination, and b) changes in when exogenous factors more or less contribute to opportunity deprivation. I.e. when an individual moves into the upper part of the lower distribution, what exogenous factors does he have or not have that made him be in that segment.

Third, as the value of α for G(α) (or ε for A(ε)) increases, the more sensitive the indicators become to distributional inequality in the upper segments of the distribution. Since only a very small proportion of the population earns or consumes disproportionately more than the national mean, the higher the sensitivity to the higher parts of the distribution means a higher value for the inequality index (i.e. see steep slope of the Lorenz curves in the upper parts of the cumulative population, or the large standard deviation for the upper quintiles). Thus, for instance, G(1) is consistently lower than G(2) no matter which aggregation methodology is used. This is the same for the Atkinson indices. Furthermore, since the variance of distributions, particularly in the upper segments for income, are higher than those for consumption, the indices estimates are relatively more apart for income than for consumption. For example, for income, the G(2) value is almost double that of the G(1), (1.110 and 0.555, respectively) and for consumption it is it is only 68% more (0.557 and 0.331, respectively). More on the difference between income and consumption inequality estimates follows under point 2.

2. Differences in results between the income and consumption inequality indices

The estimates of the inequality indices for income are higher than those for consumption, especially when comparing the two income and consumption reference variables. This is consistent with the literature and as proposed by the literature review on the measure of welfare. When computing a single-digit inequality measure, one would expect inequality levels to be lower for consumption than for income as elaborated in Section 2.2. For instance, the Gini coefficients for income and consumption are 0.541 and 0.417, respectively. Looking at the empirical literature for inequality measures and particularly Gini coefficients, it is important to highlight that, in general, the Gini coefficient does not vary much in the short- and medium-term (e.g. Salverda et al., 2009; WDI, 2010 and 2011). Minor changes in its value are usually due to significant structural and distributional changes in a society. So the relatively important difference between the income and consumption estimates in this dissertation are quite striking. Furthermore, the amplitude of the computed Gini coefficients is relatively high. When using consumption as outcome, Nepal ranks similarly to El Salvador, Gambia, Paraguay, Malawi, China, and the US (WDI for 2010). Given the nature of the NLSS III dataset (e.g. lack in fiscal information) and the developing country context, the inequality measures using consumption as economic outcome variable are preferred.

3. Differences that the various aggregation methodologies have on the inequality indices

The methodology used to aggregate any economic outcome variable does not just affect the overall value of the computed variable but also the distribution of it across HHs and individuals, and any further statistical analysis, particularly the analysis of inequality. For instance, referring to the consumption aggregates following various aggregation methodologies, Table 2.24 displays a significant variation in the results for inequality measures. Let us now focus on the Mean Log Deviation (MLD) since it serves as the key inequality in outcome measure for the subsequent inequality of opportunity analysis, as explained in Section 2.2. For consumption, the MLD results vary from as low as 0.259 when using the CBS' (2011) proposed aggregation methodology, to as high as 0.577 when following Deaton and Zaidi's (2002) past week criterion aggregation methodology. The proposed reference variable by the author, *pccons*, gives an MLD value of 0.283. This discrepancy between the various MLD results shows clearly that choosing the most adequate and individually adapted aggregation methodology to the dataset is important, because it then has a direct impact on the econometric results.

Table 2.24: Selection of Inequality in Outcome Indices using the Various Income and Consumption Expenditure Aggregates

	Generalized Entropy indices G(α) (where α is the income difference sensitivity parameter, MLD = Mean Log Deviation, CV = Coefficient of Variation)					**Atkinson indices** (A(ε), where ε > 0 is the inequality aversion parameter)		
	G(-1)	**G(0) (MLD)**	G(1) (Theil)	G(2) (half the square of CV)	Gini	A(0.5)	A(1)	A(2)
Per capita income								
Per capita disposable income (pcincome)	1.102	0.557	0.564	1.129	0.545	0.244	0.427	0.688
	(0.037)	(0.006)	(0.009)	(0.063)	(0.002)	(0.003)	(0.003)	(0.007)
Per capita income (pcincome1)	1.103	0.558	0.565	1.129	0.545	0.244	0.427	0.688
	(0.039)	(0.005)	(0.010)	(0.072)	(0.003)	(0.003)	(0.003)	(0.007)
Per capita income including utilities (pcincome2)	1.587	0.544	0.548	1.084	0.538	0.238	0.419	0.760
	(0.158)	(0.005)	(0.011)	(0.062)	(0.002)	(0.002)	(0.003)	(0.019)
Per capita income excluding wage income (pcincome3)	1.827	0.535	0.607	1.297	0.550	0.248	0.414	0.785
	(0.446)	(0.006)	(0.011)	(0.068)	(0.003)	(0.003)	(0.004)	(0.034)
Per capita income with CBS (2011) outlier constraint (pcincome4)	0.790	0.475	0.439	0.580	0.507	0.206	0.378	0.613
	(0.007)	(0.003)	(0.003)	(0.006)	(0.002)	(0.001)	(0.002)	(0.002)
Per capita consumption expenditure								
Per capita consumption expenditure including education and health, typical month (pccons)	0.316	**0.283**	0.331	0.557	0.417	0.142	0.247	0.387
	(0.004)	**(0.003)**	(0.005)	(0.016)	(0.002)	(0.002)	(0.002)	(0.003)
Per capita consumption expenditure including education and health, weekly (pcconsWEEK)	0.485	0.409	0.488	0.950	0.496	0.201	0.336	0.492
	(0.006)	(0.005)	(0.008)	(0.030)	(0.003)	(0.002)	(0.003)	(0.003)
Per capita consumption expenditure excluding education and health, typical month (pccons2)	0.299	0.274	0.329	0.578	0.410	0.139	0.240	0.374
	(0.004)	(0.003)	(0.006)	(0.016)	(0.002)	(0.002)	(0.003)	(0.003)
Per capita consumption expenditure excluding education and health, weekly (pcconsWEEK2)	0.471	0.411	0.513	1.091	0.498	0.206	0.337	0.485
	(0.007)	(0.005)	(0.008)	(0.042)	(0.003)	(0.003)	(0.004)	(0.003)
Per capita consumption expenditure, typical month, Deaton and Zaidi (2002), (pcconsDZ)	0.409	0.386	0.758	16.367	0.475	0.217	0.320	0.450
	(0.028)	(0.030)	(0.119)	(6.462)	(0.015)	(0.022)	(0.020)	(0.019)
Per capita consumption expenditure, weekly, Deaton and Zaidi (2002), (pcconsWEEKDZ)	0.683	0.577	1.156	34.403	0.575	0.309	0.438	0.577
	(0.056)	(0.045)	(0.180)	(12.370)	(0.018)	(0.030)	(0.025)	(0.020)
Per capita consumption expenditure, CBS (2011), (pcconsNLSS)	0.293	0.259	0.289	0.427	0.400	0.128	0.228	0.369
	(0.003)	(0.002)	(0.003)	(0.009)	(0.002)	(0.001)	(0.002)	(0.003)
Per capita consumption expenditure, Furuta (2016), (pcconsFURUTA)	0.419	0.398	0.796	18.078	0.482	0.224	0.328	0.456
	(0.026)	(0.029)	(0.133)	(7.392)	(0.018)	(0.020)	(0.021)	(0.018)

Source: Author's calculations. *Note*: Bootstrapped standard errors with 100 reps. in brackets.

2.7. Policy Recommendations for Inequality in Outcome

The theoretical and empirical literature on policy formulation and execution strategies to combat inequality in outcome is extensive (e.g. Bourguignon et al., 2007; Ray, 1998). This section aims to only briefly outline some of the standard inequality in outcome policies for the sake of completeness for the reader, rather than trying to execute a thorough policy framing and surveyed impact evaluation analysis. This is because, first, it is outside the scope of this dissertation to summarise this. Second, the analysis of inequality in outcome in Nepal in this dissertation is too limited in terms of data availability and statistical rigour in order to make any national or regional policy recommendations. Third, the underlying norms and judgments for any inequality policy have to be

considered an acquis by the political elites and democratic society in order to allow for effective policy implementation. Nepal has undergone continued political turmoil over the past two decades and its democratic institutions are still fragile and at an infant stage, not having democratically coined such underlying norms and judgments yet. For these main reasons, this section only briefly outlines some general inequality policy guidelines.

Traditionally, there are three main areas for public intervention for income (or economic welfare) redistribution. Their execution, however, has often proved to be challenging, as much literature has shown (e.g. Atkinson, 1975; Bourguignon et al., 2007; Esping-Andersen and Myles, 2009; Fields, 1980; Papanek, 1975; Kondor, 1975; Ray, 1998).

1. Progressive tax system (e.g. Atkinson, 1975; Bourguignon et al., 2007)

Nepal has a progressive tax system. However, it faces several challenges. The World Bank (2016, p. 4) states that "addressing inequality through ex-post redistributive policies is difficult, particularly at Nepal's level of development". A) It is difficult for the political institutions to collect taxes, avoid fraud, and individual leakage through the system. For instance, one key strategic pillar for Nepal's economic development is remittances from labour migrants (ca. 25% of GDP; Nepal Ministry of Finance, 2011; ADB, 2018). The government here faces the trade-off between encouraging individuals to seek economic activity abroad in order to remit money back into Nepal, and upon returning, contribute through brain gain to the aggregate economic performance, and in taxing those remittances. Anyhow, most of these remittances leak through the system and remain unreported (informal exchange with employees from the CBS, Fall 2018).

B) The democratic institutions in Nepal are still at an infant stage, so they lack both in know-how and credibility amongst its citizens when it comes to tax collection.

C) Nepal is one of the 30 most underdeveloped countries in the world (WDI, 2011), where more than 85% of the population live below 2USD per capita per day when using consumption expenditure as the economic outcome variable (see Table 3.27: Most Opportunity Deprived Types and Their Profile for Per Capita Income). When the entire population is so impoverished, it makes it difficult to impose further financial burdens on individuals. Furthermore, many regions in the country are so remotely rural, and the people so poor, that the incurred cost for collecting, for instance, 50 cents per capita per year is much higher than the actual tax.

D) Another pillar of the political development strategy for the country is international aid. This makes up 25% of national GDP (ADB, 2018). Despite all the benefits of it, it undermines the mutual responsibility between citizens and political institutions and creates a vicious downward spiral for economic dependence on international aid. During the author's field trip to Nepal, he exchanged formally and informally with representatives from various governmental, non-governmental, international governmental, and international organisation institutions. A consensus was reached during the informal discussions that Nepal's decade-long dependence on international aid, both private and public, as well as foreign policy interests (particularly with the People's Republic of China and the Republic of India competing for Nepal's natural resources), has impeded the country to develop a dynamic and multi-layered economy. Consequently, the corporate tax contributions for Nepal's public budget are minimal to where they could be.

These are some of the reasons as to why Nepal is challenged in combatting inequality through a progressive tax system. Nepal should focus on first solidifying its political institutions and constructing the basis for a competitive and multi-layered economy. Yet, as of the fiscal year 2018/2019, the Nepalese government has introduced reforms to the progressive tax system by trying to simplify it and easing the burden on low- and middle-income earners and increasing the proportion of taxes paid by high-income earners. For instance, individuals earning up to 350,000 NRs. only have

to pay 1% tax, while individuals earning between 650,000NRs. and 2,000,000NRs. are charged 30% (Dhungel, 2018).

2. Land reforms

Land reforms have been a popular means for combatting inequality particularly, on the Latin American continent, yet the effectiveness is highly debated both in academic research and at the public opinion level (e.g. Foxley, 1976; Barraclough and Domike, 1966; Ray, 1998; Taylor and Hudson, 1972). The Nepalese government has not undergone any significant efforts to implement land reforms. This is largely because most valuable land is held by the upper castes, who still play the most significant roles in the government. Also, the Nepalese tax system does not have extensive wealth and property tax. Yet, the new 2015 Constitution, which serves as the basis for the federalisation process, aims at giving significantly more powers to local authorities, particularly when it comes to tax (and wealth tax) collection (Nepal Law Commission, 2015). Time will tell whether the implementation of this will succeed.

3. Social redistribution schemes (e.g. minimum salary, social welfare benefits)

Nepal is yet an infant democracy, continuously trying to shape its institutions and societal political engagements. Furthermore, it has extremely scarce public resources. So, Nepal is yet far from becoming a social welfare state with effective social policies, such as universal health care access, minimum wages, or unemployment benefits, all of which would help to alleviate inequality (Ray, 1998). Because Nepal is not yet able to provide basic social welfare and security, individuals are forced to rely on the traditional and conservative safety nets, namely the caste system, religious belonging and groupings, and social capital, and so on, forcing people to place from their safety nets first. This in turn leads to direct and indirect discrimination of people outside of one's group. Yet, Nepal has already started to write laws, which are amongst the most progressive in the region, to fight vigorously against discrimination of any kind. As an example, Nepal has equalised human rights for all people in spite of their ethnic or caste belonging, religious beliefs, gender, sexual orientation, etc. Furthermore, during the current ongoing federalisation process, the government is trying to implement minority representation quotas in its public institutions. Whether the execution of these will be successful is debatable, but at least these are some measures the government tries to take in order to fight discrimination and allow the poorest of the poor to make a living for themselves.

The above suggested measures are some that have been historically employed and debated to combat inequality in outcome. Much of their effectiveness, however, depends on the strength of the political intuitions and the implementation of the policies. Since 2016, Nepal has been undergoing political restructuralisation, trying to become a federal state. This process is expected to take more time and still much political uncertainty remains.

Furthermore, Nepal has to undergo thorough political debates on its development strategies. Bouguignon et al. (2007) discusses the trade-offs between poverty alleviation, inequality mitigation and economic growth policies. Additionally, any inequality policies should reflect the ethical and social norms of the society in order to guarantee societal and political stability (e.g Bourguignon et al, 2007; Kondor, 1975). This is all the more important in economically, socially, and politically fragile Nepal. A serious, progressive, and fact-based debate on this must lay the foundations for long-term inequality and development policies. Once these fundamental questions are answered, holding a democratic majority, the roots of inequality must be isolated and analysed in order to formulate the

most adequate and individually tailored inequality policies. Pinpointing the multidimensional concept of inequality is easier said than done, and as argued in Section 2.2 and the following chapter on inequality of opportunity, inequality is more than just the unequal distribution and disparity of income.

2.8. Conclusion and Literature Contributions: Inequality in Outcome in Nepal

The purpose of Chapter 2 was threefold. First, to compute the **most suitable economic welfare variables using the most suitable aggregation methodology**. Consumption expenditure and income were identified as the most commonly used economic welfare proxies, both bearing their advantages and disadvantages. Of the two, consumption expenditure is the preferred measure given the Nepalese context. This is because it is seen to more adequately represent the current living standards of the people than income. Also, income bears a temporal component (current and future welfare) and, therefore, is arguably less appropriate for this dissertation's subsequent inequality of opportunity analyses. It was highlighted that great care must be taken when aggregating the economic welfare variables. Numerous guidelines exist to do so, and the estimated value of the aggregates vary depending on which of them are followed. The author computed consumption and income aggregates following various guideless and then drew conclusions on which elements in the aggregation process of them were the most suitable given the Nepalese context and the nature of the NLSS III dataset. So, one "ideal" consumption and one "ideal" income aggregate are proposed using an "optimised" aggregation methodology. It is crucial to understand how the economic outcome variables were computed when interpreting the consequent descriptive distribution disparity and the further statistical econometric results for both the inequality in outcome indices and the inequality of opportunity analysis in the next chapter. In other words, different aggregation methodologies have different impacts on the statistical analysis and their results.

The second purpose of this chapter was to **describe the distribution and disparity of economic outcome** across Nepal using both graphical illustrations and descriptive statistics based on the identified aggregated economic outcome variables. To further complement the nation-wide disparity description, distributional inequalities were broken down by population quintiles and by geographic zones. A large degree of disparity was noted for all outcome variables at the national level and at the geographical breakdown level. However, the degree of disparity depends to some extent also on the aggregation methodology used to compute the outcome variable.

The third purpose of this chapter was to use identify **the most universally recognised inequality in outcome indices**, weigh their pros and cons, and to estimate and describe them. The most commonly used Generalised Entropy and Atkinson Inequality Indices were computed. Different inequality in outcome indices place emphasis on different parts of the distribution, thus giving different results. Judging from the literature review and the inequality in outcome analysis, the Mean Logarithmic Deviation (one of the Generalised Entropy Inequality Indices) was identified as the most suitable inequality in outcome reference variable for the subsequent inequality of opportunity analyses. Two of its key advantages are that it is the only measure of the GE indices to satisfy the path independence axiom and that it is decomposable between income sources and between individuals.

To conclude on the **entire Chapter 2** it mainly described the distribution, disparity and inequality of income and consumption in Nepal. However, it did not manage to explain these with rigour. It is difficult to generalise on the fairness and social justice of a society using a purely descriptive disparity and single-digit inequality in outcome measures. Inequality is a multifaceted and complex concept, and in order to explain it, more complete measures of inequality need to be applied.

The next chapter proposes a measure of inequality that seeks to complement the inequality in outcome analysis. It seeks to explain and quantify the factors that most impact on the inequality in outcome. To do so, a measure for inequality of opportunity is proposed, which aims to quantify, in a first step, the proportion of inequality in outcome that is due to opportunity deprivation. In a second step, it decomposes this inequality of opportunity in terms of explanatory factors. The inequality of opportunity analyses are done at the national level, but are also decomposed by geographical zones and by population groupings.

There are some, albeit minor, **contributions to the literature of economic welfare** measurement and inequality in outcome. First, this chapter summarised and compared the literature on how to compute the income and consumption aggregates, suggesting adaptations to the NLSS III dataset and the Nepalese context. Given that there is little consensus on how to compute consumption, especially, and to a lesser extent income, this comparison can be a valuable reference for researchers interested in thoroughly looking through the aggregation methodology of economic outcome.

Second, this chapter highlighted the importance of paying attention to correctly aggregating income and consumption and knowing the exact aggregation procedures. This is because the distribution and disparity of the economic outcome variables are affected differently by different aggregations methodologies, which in turn affect any further econometric inequality analysis and decomposition. Also, knowing the aggregation procedure of the dependant variables is crucial for interpreting the econometric results and drawing any policy recommendations.

Last, an extensive descriptive analysis of the distribution and disparity for economic outcomes in Nepal was done. This was complemented with the computations for the most commonly used inequality indices. The results were cross-compared and a brief attempt was made to explain them. This detailed description of economic outcome distribution, disparity, and inequality can serve as a first reference for future inequality research on Nepal.

2.9. Need for a More Complete Inequality Analysis: Inequality of Opportunity

The level of inequality in economic welfare can be analysed in many ways that are complementary to each other. First, inequality can be analysed in a descriptive manner by looking at the distribution of income across the country by population, and by geographical dispersion, for instance. Second, a more sophisticated statistical analysis can be made drawing on the distributional disparity of economic welfare using, for instance, the Gini or Theil indices. They can provide additional information using the distributional characteristics for understanding the disparity in economic welfare across a society. These two steps consider that the underlying concept of inequality is the uneven or "unfair" distribution of economic welfare across a society using, for example, income as a monetary measure of welfare. However, one can argue that not only income and consumption, but also other elements such as life expectancy, health status, or education level, contribute to one's welfare or living standard. Just like the concept of welfare, the concept of inequality is also multifaceted, and one must move beyond a unilateral monetary measure to a more complete measure and analysis of it. So, while the analysis of economic disparity is crucial in serving as a starting point for inequality analysis, other inequalities must be considered. These include, for instance, educational and health inequality, nutritional inequality, social mobility inequality, or various market access inequalities (e.g. labour or credit market).

The next chapter tries to build on the inequality literature, and more precisely on a relatively recent sub-branch of it: inequality of opportunity. Inequality in opportunities is the core thought behind development economics. Most topics in development economics refer to the key concept of opportunities and access to them; such as credit market access, the inequality-poverty-economic

growth triangle or child labour. Yet, the term inequality of opportunity only became a reference with Roemer's 1998 benchmark book establishing the field of opportunity economics. Other concepts, such as intergenerational (e.g. income or educational) mobility, inequality decomposition or just inequality, could be perceived as closely related but distinct from the concept of inequality of opportunity. "Opportunity economists" (e.g. Bourguignon et al., 2007; Checci and Peragine, 2010) argue that these terminologies are partially incoherent because they include both static and dynamic elements as well as inequalities in outcome and opportunity that arise due to individual choice and exogenous factors.

The literature on inequality in outcome (e.g. income) is abundant. It may, however, be argued that the distribution of particular outcomes is not appropriate for assessing the fairness of a social system or an allocation (e.g. Sen 1985, Arneson 1989). Inequality that arises due to differences in choice (attributes for which individuals can be held responsible and accountable) and, therefore, different outcomes, is more ethically acceptable. In other words, income gaps that arise from differences in efforts are generally less objectionable than those that are due to differences in parental background or gender for example. The goal of the decomposition of inequality of opportunity is to single out how much inequality is due to unequal *opportunity*. Thus, the next chapter tries to complement the inequality in outcome analysis of Nepal, with an in-depth analysis of inequalities of opportunities.

Chapter 3 Inequality of Opportunity as an Alternative Inequality Measure: Empirical Analyses of Nepal

In order to fully analyse the existence of IOp in Nepal, this dissertation consists of two main analytical chapters. Chapter 2 investigated the level of inequality in outcome (income and consumption expenditure), which serves as the basis for the subsequent analysis of IOp in Nepal in this Chapter (Chapter 3).

This chapter, first, gives an extensive discussion of the IOp literature, its terminologies, basic frameworks, underlying principles, challenges and trade-offs in computing the IOp indicators, and interest for the public and policy makers to study IOp. Next, the methodology is explained in both an intuitive and a theoretical manner and the country-specific background is given. The dataset is described in more detail and the applied dataset constraints, the author's attempts to overcome some of the dataset imperfections, the choice of the dependent and independent variables, and the econometric model adaptations to the Nepalese dataset are discussed and justified. Four separate analysis sections follow for the study of IOp in Nepal: a national level, Development Region level, urban rural area level and population grouping by economic welfare quarter level. Each of these provides a descriptive and econometric analysis containing the computed scalar IOp indices, RF-OLS regressions, specific circumstantial shares of unequal opportunities, identification of the profiles of the most disadvantaged types and various robustness checks. Then the main results from all four levels of analyses are compared. Some policy implications, study limitations and further extension ideas are given. Finally, Chapter 3 concludes.

On a side note, Chapter 2 and 3 are complementary for the wholistic understanding of IOp in Nepal. Chapter 3, however, is the main focus of this dissertation, thus its disproportionate length compared to the other chapters. While Chapter 3 could be divided into two chapters (national level analysis and disaggregate level analysis), due to a large overlap in terms of methodology, results, policy conclusions and so on, it was decided to keep the two chapters as one.[4]

Chapter 4 compares and highlights the literature contributions of the previous two chapters on a methodological and empirical level. Chapter 5 concludes this dissertation.

3.1. Literature Review

This section discusses the IOp literature. To do so, first different terminologies of *inequality* are reviewed. Then, the basic concept of IOp and its definition by Roemer (1998) are described. Furthermore, some additional underlying theoretical and ideological principles are explained in order to fully understand the complexity of the IOp concept. Next, some of the key IOp conceptual challenges and trade-offs are discussed. Then, the importance of studying IOp is highlighted by looking at changes in public attitudes, the underlying normative theory and its significance for policy makers. Following this, the main econometric IOp frameworks are explained in order to stress the advantages and disadvantages of the framework applied in this dissertation. Finally the literature review concludes.

[4] The aim is to publish Chapter 3 as two papers. First, the national level IOp analysis with an emphasis on comparison to the existing literature, most of which also only focuses on national level IOp computation. Second, the disaggregated national level IOp analyses adding to the IOp literature in terms of depth and thorough understanding of IOp with its regional and population differences across one country. This then allows for more customised and effective IOp policy framing. The second paper would thus be a more Nepal contextual specific analysis with important policy implications.

The section after the literature review focuses on the empirical IOp model used in this study, which is largely inspired by Ferreira and Gignoux (2011; from now FG, 2011).

3.1.1. Inequality of Opportunity: Terminology

What is inequality? Is all inequality bad? And how is it measured?

In order to understand what the term IOp means, it is essential to briefly elaborate on other terminologies that could be perceived as, but are distinct from, IOp. First, the term inequality and some basic measures of inequality are elaborated. Section 2.2 of this dissertation already elaborated extensively on the traditional perceptions and most common measures of *inequality in outcome*. So this section builds on what was said before. Second, the terms *intergeneration (income) mobility* and *meritocracy* are explained. Finally, the term *inequality of opportunity* and some largely synonymous expressions are examined.

It is important to bear in mind that these terms and their corresponding various measures have differences in temporal (static versus dynamic) perspectives and input versus outcome elements.

The topic of inequality has always been at the heart of development studies. When the question arises of what development actually means, the key elements referred to generally include economic growth, poverty or inequality. The inequality-poverty-economic growth triangle has long been a source of fascination and contradicting research (Bourguignon et al., 2007). Little consensus exists on the directions of causality in the empirical research. In other words, what causes what, in which way and to what extent. Historically, the interest in the study of inequality is mainly due to ethical but also practical reasons, where, for example, greater or less inequality is thought to impact differently on economic growth.

It is difficult to define ***inequality*** because the term is so generic and multifaceted. That is why there is such an array of definitions and measures for inequality. According to Ray (1998), the most basic and often referred to measure of inequality is per capita income inequality (i.e. ***inequality in outcome***). It is relatively easy to quantify since it is based on unequal distributions of easily measurable income. As seen in Section 2.2, the list of inequality in outcome measures is endless, each placing emphasis on different parts of the distribution. Also, the section distinguished between the terms ***distribution***, ***disparity*** and ***inequality***.

Barros et al. (2009) summarise that until the 1970s, the common understanding of assessing fairness and equity of social allocation was primarily based on the distribution of outcomes (i.e. income or consumption expenditure). Rawls (1971) was one of the first to question the until then narrow views of fairness with his book on the "Theory of Justice", introducing elements of opportunity, liberty and primary goods that allow to level playing fields in Roemer's terms. Since then, much debate and literature have emerged on the *Theory of Distributive Justice* (e.g. Rawls, 1971; Sen, 1985; Roemer 1993; Cohen, 1989; Salverda et al, 2009; Barry, 2005) shaping public opinions of the perception of inequality and what is considered as *fair* or *unfair*, or *legitimate* or *illegitimate* inequality.

As a reminder, the ***Mean Logarithmic Deviation*** **(MLD)** of the Generalized Entropy (GE) inequality measures is chosen to serve as the most suitable inequality in outcome reference variable for the IOp analyses in this chapter. It is computed for both per capita income and per capita consumption expenditure. Reasons for this were elaborated at great length theoretically and empirically in Chapter 2. To put it simply, the MLD is the percentage difference between the economic outcome of a randomly selected individual and that of the sample average (Haughton, 2009). More formally, it is the logarithmic value of the outcome distribution, and it is particularly sensitive to

distribution inequalities in the lower and middle segments of the distribution. Some of its key advantages are that it is the only measure of the GE indices to follow the path independence property and that it is decomposable between individuals, groups and sub-groups.

Intergenerational mobility (IGM) is a frequently used term that can sometimes be confused with the term IOp. Intergenerational mobility generally studies the impact of parents' socio-economic status (e.g. earnings, education attainment) on that of their children and how this evolves over time. Much literature on this exists (e.g. Fields, 1980; Morgan et al., 2006; Salverda et al., 2009, Ch. 21). There are numerous studies on IGM and a particularly great amount in the field of labour economics and income earnings and are applied to the US (e.g Van de Gaer et al. 2001; Solon, 1999,). Corak (2013), for example, studies intergeneration mobility and finds that particularly family background and education play a significant role in labour market earning immobility. Furthermore, Solon (1999), for example, uses the popular methodology of twin studies in the US to test for intergenerational mobility. Complete mobility is said to be achieved when outcome in time 2 is totally independent to time 1. More mobility therefore means more equality. The difficulty however is how to measure and define mobility. The authors above use transition matrices (e.g. employment position, education attainment) for this.

There are a few overlaps between the concept of IOp and IGM (e.g. Barros et al., 2009). The first is the inborn, genetic and exogenous elements of ability and skills that are passed on from one to another generation and which impacts on outcomes to some extent. The second is the relevance of efforts (in the IOp literature) and social capital and one's own ambitions (in the IGM literature) in influencing the status of the socio-economic status in question. Barros et al. (2009) argue that IGM result reliability is scarce, as can also be argued for IOp estimates. A third element is the overlap in policy recommendations. Both, for instance, recommend education as a tool for equalising opportunities and enhancing mobility.

There are however a few elements that distinguish the IGM and IOp concepts. First, IGM is a dynamic or temporal measure of inequality (Dearden et al., 1997), while IOp is mostly a static measure. Second, equalising opportunities does not imminently lead to greater economic mobility. Equality of opportunity policies (EOPs) cannot alleviate the impact of inherited ability, luck or ideals that explain an important part of outcome inequalities and economic mobility (Barros et al., 2009). IOp is however employs a much richer set of explanatory variables in trying to quantify an overall or partial impact of these on outcomes.

So, the IGM approach can be seen as similar, but yet distinct to the concept of IOp.

The term ***meritocracy*** is sometimes seen as synonymous to IOp. Indeed there are some overlaps to the extent that individuals should be rewarded for their efforts, skills and performance. However, a clear distinction must be made. Roemer (1998) emphasises the importance to level-off playing fields (in other words, give everyone an equal set of tools through education, for instance) before individuals are left to themselves (own responsibility and accountability) and compete for positions within society. From that stage on, the concept of IOp equals that of meritocracy, where individuals who compete, should solely be based on the attributes of their effort rather than socio-economic exogenous elements (Barros et al., 2009). Meritocracy does not discriminate on circumstances, but only on efforts, skill (i.e. performance) and "produced social outcomes", while the IOp concept only takes into consideration efforts (and not performance) and the fair treatment of all.

The term ***IOp*** is not often used, because it includes static and dynamic, exogenous and endogenous elements. The term IOp has great appeal due to its generic nature especially in politics, yet it is hard to pin point and *opportunity* is hard to measure. Therefore, it is less risky to observe mobility and measure inequality in terms of outcome. The literature on the measure on inequality in outcome (e.g. income, consumption expenditure) is abundant. It may, however, be argued that the

distribution of particular outcomes is not appropriate for assessing the fairness of a social system or an allocation (e.g. Sen 1985, Arneson 1989). Inequality that arises due to differences in choice (attributes for which individuals can be held responsible and accountable) and therefore differences in outcome, are more acceptable. In other words, income gaps that arise from differences in efforts are generally less objectionable than those that are due to differences in ethnic or gender discrimination for example. There is relatively little literature on circumstance and effort inequality, but predetermined circumstances have become central to the literature on social justice in political philosophy, social choice, and increasingly in mainstream economics. The goal of the decomposition of IOp is to find out how much inequality is due to *opportunity* or *opportunity deprivation*.

Roemer (1998) set a theoretical benchmark for the field of opportunity economics by dividing IOp into efforts (factors an individual can control, endogenous), circumstances (factors an individual cannot control, exogenous) and advantages (economic outcomes such as income). He defines *equality of opportunity* (from now on EOp) as a situation in which outcomes are independently distributed from circumstances. This means that when perfect EOp is observed, differences in outcome are only due to differences in effort or choice (up to one's personal responsibility). Therefore, it appears ethically acceptable to have inequality in outcome. This inequality is more socially acceptable than inequality in circumstance, factors one cannot change. Often this is referred to as *fair* or *legitimate* inequality, while inequality arising from differences in circumstances are often referred to as *unfair* or *illegitimate* inequality.

The study of IOp allows measuring the overall and partial, direct and indirect effects of circumstances on economic advantage. It tries to quantify the overall impact of exogenous variables on differences in economic outcome as well as singling out the significance of each circumstance variable. FG (2011) call IOp the *true* measure of inequality.

Furthermore, the difference between ***inequality of opportunity*** and ***discrimination*** may be questioned. Most of the distinction lies in the etymology of the vocabulary[5]. In sum, the terms ***opportunity deprivation*** and ***discrimination*** can be used as synonymous, to the extent that both refer to exogenous circumstances. For instance, "Muslim women in Nepal are opportunity deprived in accessing the labour market" can be seen as synonymous to "Muslim women in Nepal are discriminated against when accessing the labour market."

3.1.2. Inequality of Opportunity: Basic Concept and Terminology (Roemer, 1998)

This section elaborates on the basic IOp concept as developed by Roemer (1993b, 1998).

---

[5] The Oxford English Dictionary (OED, online) defines ***discrimination*** as " the treatment of goods, trading partners, etc., on a more or less favourable basis according to circumstances; an instance of this" and as an "unjust or prejudicial treatment of a person or group, esp. on the grounds of race, gender, sexual orientation, etc.; frequently with against. Also (with in favour of): favourable treatment of a person or group, in order to compensate for disadvantage or lack of privilege." The OED states that discrimination is an action word, originating from the action verb "to discriminate against something or someone" in the sense of "making a distinction". The term inequality of opportunity is a composite noun and cannot be used as an action noun. Instead, opportunity deprivation can be employed.
The OED defines ***deprivation*** as "the action of depriving or fact of being deprived; the taking away of anything enjoyed; dispossession, loss." In this case, the act of depriving someone or groups of people from opportunities.
The OED defines ***opportunity*** as "a time, condition, or set of circumstances permitting or favourable to a particular action or purpose."

Roemer launched a philosophical and ideological debate about the concept of IOp with his benchmark book entitled "Equality of Opportunity" in 1998. Consequently, his IOp framework has become the initial point of reference for everyone eager to do research on the topic of opportunities in economics.

According to him, *opportunity* may be defined as "the chance of obtaining a good or service if one seeks it" and *equal opportunity* means that individuals may be held *responsible* and *accountable*[6] for the "anticipated consequences of their voluntary actions". We should therefore not care about the outcome of choices, since they are our own responsibility and we are accountable for them, but one should study what the reasons are behind the taken decisions. Basic microeconomics indicates that different utility functions give different choices and therefore individuals' trade-offs between consumption and leisure, for instance, exist (Varian, 2010).

While the definition of opportunity sees a large consensus, the concept of *equity* does not. What is just? What is fair? Is inequality fair or unfair? – This is largely a normative and ideological debate, out of the scope of this dissertation. Section 3.1.4, however, briefly expands on the "efficiency versus equity" and the "normative versus positive" trade-offs. Yet, Roemer (1998) describes three main economic visions of equity and equality, the first two being *equality of welfare* and *equality of resources*. This study, however, focuses on the third vision, *equality of opportunity (EOp)*. **Equality of welfare** may be considered as a poor ideal for numerous reasons. For example, the rationality and the weighting of preferences may be questioned. Also the ideal of full information is problematic. These challenges bring about, amongst others, the *Problem of Slavery of the Talented* and the *Social Biological Problem* (Roemer, 1998). Arneson (1989) gives a complete debate on these two problems. **Equality in resources** has been debated largely and particularly in economic literature of communism and socialism and is not explained here. Later, arguments are given as to why some believe that **equality of opportunity** is the best concept for the equality ideal.

Roemer (1998) elaborates two views of EOp. First, the *non-discriminant principle*, which means that everyone is in competition for a position in society (e.g. employment position) and individuals should only be judged on attributes of their performance and not on exogenous socio-economic factors such as race or gender. Second, society must do everything possible in order to *level playing fields* among persons who compete for positions. Levelling-off playing fields should be done especially during the educative years, so that all those who have potential attributes can be considered. From this, one can conclude that there is a *before* and an *after* competing for a position. Before the competition starts, opportunities should be equalised (if necessary, by public intervention) and after, individuals are left on their own. One can interpret this as equalising opportunities by giving everyone an equal set of tools (i.e. education), but what one does with these tools is up to everyone's own choice and responsibility. Roemer (1998) calls IOp intervention *Equality of Opportunity Policy (EOP)*.

The next normative debate is as to where to draw the borderline between the before and after. Roemer (1998) tries to determine this by making a clear distinction between so-called inequality characteristics and source of inequality characteristics. To do so, these characteristics are divided into two main categories. First, the endogenous information and decisions variables he calls ***effort*** (such as how much time is spent studying, personal ambition, social capital formation, decisions taken or aspiration). The term effort is often also referred to as *responsibility* or *choice*. Second, the exogenous variables that are out of personal responsibility and accountability called ***circumstances*** (i.e.

[6] Roemer (1993a) has an elaborate debate about the differences in individual's accountability and responsibility. This, however, is a more normative than economic debate, thus not elaborated here.

exogenous socio-economic factors such as parental education, ethnicity or gender). Individuals are categorised into different subgroups of people who are homogenous in their circumstances. Roemer (1998) refers to these subgroups as ***types***. This means, that any inequalities in outcome that should arise within and between types would be due to effort and not circumstance differences. This situation would be referred to as perfect *Equality of Opportunity* (EOp).

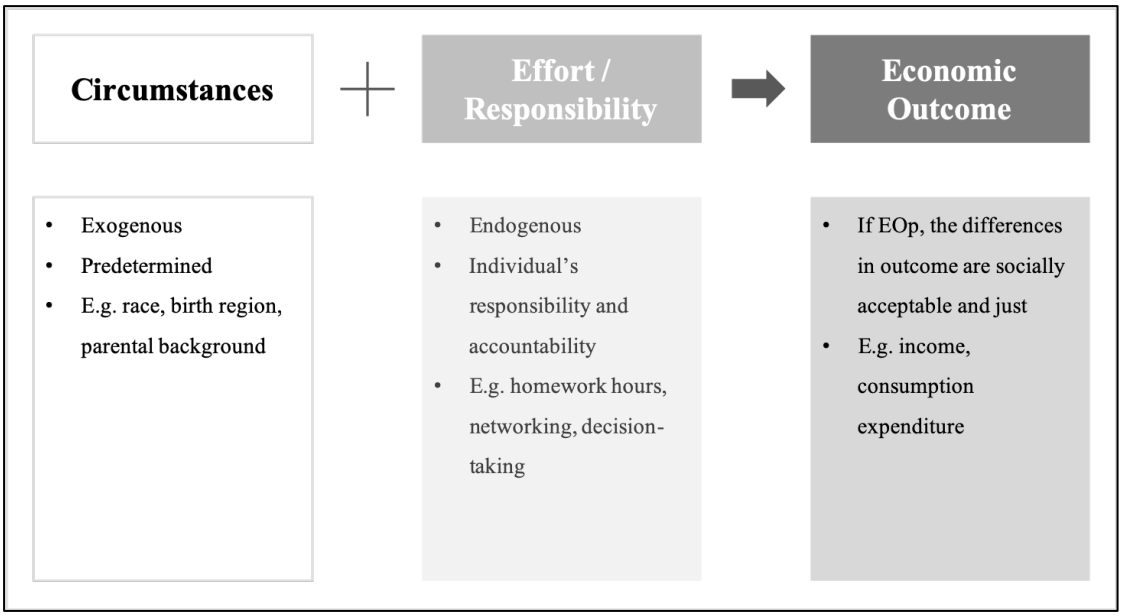
The following quote by Roemer (1998) beautifully summarises his entire IOp concept. It contains all the previously discussed elements.

> *"What society owes its members, under an Equality of Opportunity Policy, is equal access; but the individual is responsible for turning that access into actual advantage by the application of efforts."*

Roemer (1998) further uses a famous quote by José Ortega y Gasset, a prominent Spanish philosopher, as his book dedication to illustrate the concept and the philosophical debate behind IOp: "Yo soy: yo y mi circumstancia." In English this translates into " I am: I and my circumstances". What Ortega y Gasset wants to say is that one needs to acknowledge that everyone has their own set of predetermined conditions which they cannot influence, but everyone retains a lot of individuality.

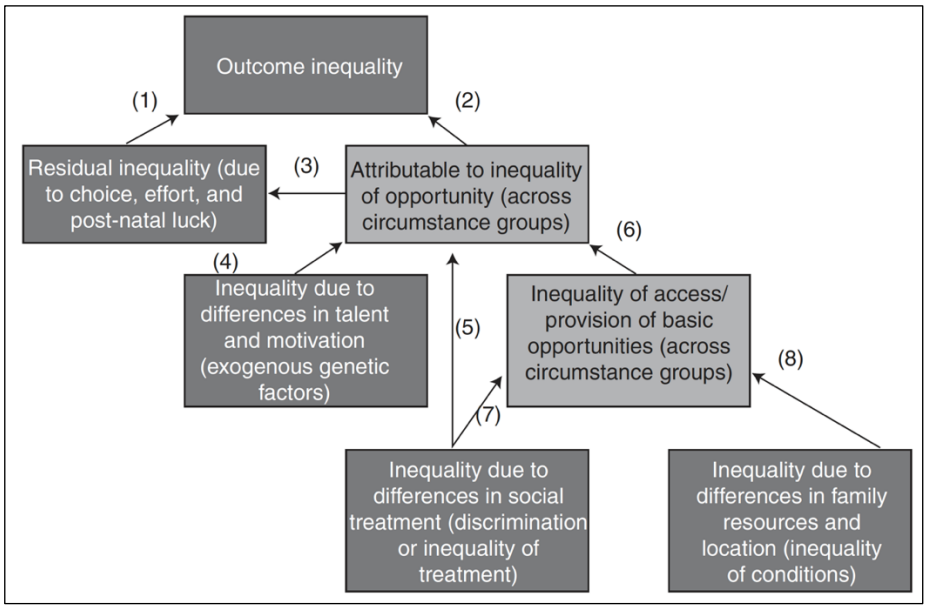
Figure 3.1 gives a simplified visual overview of Roemer's (1998) IOp concept. **Figure 3.2** by Barros et al. (2009) complements Figure 3.1 with more details of the decomposition of outcome inequality. The authors agree that outcome inequality arises primarily by two factors: efforts and circumstances. However, circumstances are further decomposed into exogenous genetic factors (e.g. IQ or talent) and circumstance variations between groups (e.g. public goods provision). The latter is then further decomposed into two circumstantial elements, opportunity deprivation based on exogenous social factors (e.g. ethnicity) and exogenous economic factors (e.g. due to parental lack of resources, born in an impoverished neighbourhood). In sum, Barros et al. decompose the circumstance variable Roemer refers to into four sub-elements.

Figure 3.1: Basic Inequality of Opportunity Concept (Roemer, 1998)



Source: Author's elaboration based on Roemer (1998)

Figure 3.2: Outcome Inequality Decomposition (Barros et al., 2009)



Source: Barros et al. (2009, p. 31)

There are various **challenges and trade-offs** with the IOp concept, which are discussed later (see Section 3.1.6.2 for a complete debate on this). However, one key challenge to mention now briefly lies in how to differentiate between effort and circumstances. For example, it is hard to differentiate whether someone attains a certain level of education due to effort, circumstances or a combination of both. Yet, the division of IOp into a) endogenous effort and b) exogenous circumstance variables and c) economic outcome variables, has set the baseline for the literature on IOp.

There are a three main clusters of **the IOp concept that make it very appealing** to scholars, the general public and policy-makers. First, according to Banerjee and Duflo (2003) the **empirical literature on inequality** is inconclusive, because there is a conflict between two kinds of inequality. When looking at the relation between inequality and economic growth, for example, firstly elements that belong to outcome inequality (e.g. effort) actually have a positive effect on economic growth. Secondly, elements that are considered to belong to unequal opportunities (e.g. access to schools, credit market) and that are considered to have a negative effect on growth. They therefore argue that the literature on inequality is inconclusive and one should separate circumstance and effort variables. Marrero and Rodriguez (2009) and Banerjee and Duflo (2003) obtain the same effects of circumstances and efforts on economic growth in their studies. An additional reason as to why the concept of IOp up until recently was not often used is because it includes both static and dynamic elements and lack of data availability.

Second, there has been a **shift in ideals** away from inequalities in outcome towards inequalities of opportunities. As Roemer argues, inequalities in outcome that arise from differences in efforts (such as decision-making, studying, building social capital) are considered more socially and ethically acceptable than inequalities arising due to differences in exogenous socio-economic conditions. In other words, this rhetoric is often associated with slogans like “the American Dream” or “self-made man”. A hope would be that instead of comparing countries’ inequality levels with the Gini coefficient (i.e. coefficient for measuring inequality in income distribution), for instance, one could use a universal IOp index. In fact, the World Bank has been working on what they refer to as the Human Opportunity Index (HOI; e.g. Barros et al., 2009). Yet the development of a global HOI is still at an infant stage. It is constraint by data availability and does not allow for a dynamic evolution picture of opportunities. Thus the World Bank is increasingly shifting to what they refer to as the Human Capital Index (HCI),

which is quite distinct from the HOI. The HCI place more emphasis on the study of individual circumstances, while the HOI tries to portray a more global opportunity deprivation picture which is then decomposable in terms of circumstance. The HCI is less demanding in terms of data than the HOI (personal discussion with WB economists who worked on the 2018 "Fair Progress Report"; Narayan and Van der Weide, 2018).

Last, the IOp concept can be a useful source to formulate **public policies**. Section 3.1.5 looks more closely into the appeal of the IOp concept for policy-makers. However, more research into the concept and its empirical application need to be done in order for the results to be reliable enough to serve as the basis of policy formulation. Yet it can help to customise and localise EOPs in order to maximise their effectiveness by targeting the most opportunity deprived groups., Roemer therefore argues that there is a **strong link between political philosophy and welfare economics**. Thus, all scholars should be literate on both issues.

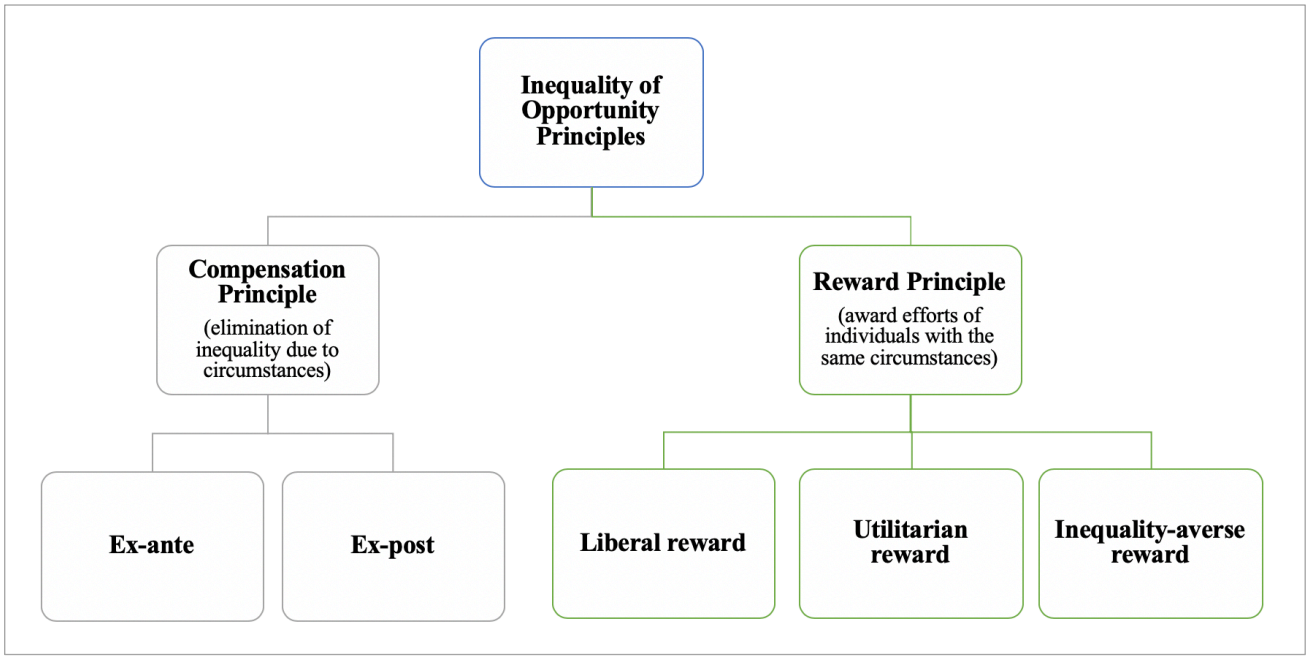
3.1.3. Inequality of Opportunity: Other Principles

The goal of this section is to paraphrase the complex and heavily debated underlying principles behind the numerous IOp methodologies. A series of literature surveys exist (e.g. Ramos and Van De Gaer, 2016; Ferreira and Peragine, 2016; Roemer and Trannoy, 2015). This section, however, follows the way of illustrating the various principles as discussed by Ramos and Van De Gaer (2012).

It is important to have at least an overview of the various underling principles for the various IOp methodologies in order to justify and continue with the methodology used in this dissertation. Only a brief overview is given since an exhaustive elaboration of all principles would be out of the scope of this dissertation, since only one of these principles serves as a baseline for the methodology employed in this chapter of the dissertation.

Ramos and Van De Gaer (2012) summarise the theoretical literature of IOp into two main principles. The first principle is referred to as "compensation principle". This principle emphasises that all inequalities arising from differences in circumstances should be eliminated In other words, individuals should be "compensated" for any differences in economic outcome that arise due to different circumstances. The study of IOp in general circulates around the analyses of the circumstances on economic outcome. The second principle is referred to as "reward principle". This principle emphasises that individuals with the same circumstances should be rewarded for their differences in effort. The study here generally circulates around the analysis of effort and that individuals should be "rewarded" for their efforts. Figure 3.3 illustrates the different principles and their main sub-categories.

Figure 3.3: Overview: Other Inequality of Opportunity Theoretical Principles
(Ramos and Van De Gaer, 2012)



Source: Author's elaboration based on the descriptions by Ramos and Van De Gaer (2016)

The **compensation principle** can be sub-categorised into A) an ex-ante and B) an ex-post approach.

A) the **ex-ante IOp approach**, like most ex-ante approaches, focuses on the prospects of outcome before any action is taken. EOp is said to exist when everyone has the same opportunities, despite differences in circumstances. For example, despite belonging to an ethnic minority or majority everyone should get the same expected outcome (i.e. same access to the same level of income). In order to compare expected to actual outcome distribution inequalities, individuals are divided into *types* (or groups) of people where they are homogenous in terms of their circumstances.

One major advantage of this approach is that efforts do not need to be observed. This is because expected outcome levels are normally measured in a way that looks at the distributional outcome of people within the same type and inequalities between types. This is what is done by the approach used in this dissertation based on the model elaborated by FG (2011). In their and in this study, the effort variable is not observed and does not need to be included since expected type outcomes are computed under the assumption that everyone has the same opportunities. More on the precise direct ex-ante model (under the compensation principle) used by FG (2011) in Section 3.1.6.

Roemer and Trannoy (2015, p.284-289) refer to the ex-ante approach used by FG (2011) as the *Direct Unfairness Approach* (DU). This is because the inequality estimate is the *unfair* inequality that arising from different circumstantial treatment (i.e. ex-ante in circumstances). They define it as "[the Direct Unfairness or ex-ante approach] is computed as the inequality of the counterfactual distribution when one has removed the effect of effort variables, either by suppressing them, or by imputing to each individual a reference value of effort such as the average value. […]. For the reduced form, a natural choice for DU is to compute the inequality of the conditional expectation of outcomes across types […] which is a neat solution chosen by Ferreira and Gignoux (2011) The residual is set to 0, its mean value."

Summarizing, the ex-ante approach focuses on expected outcome. EOp is said to exist when everyone has the same opportunities, despite a differences in circumstances. One major advantage of this approach is that efforts do not need to be observed (or measured). This is because expected or

hypothetical outcome levels are computed. In other words, the ex-ante approach measures the inequality of the average outcome across types.

B) The compensation principle's **ex-post approach** focuses on differences in outcome between individuals with the same effort but different circumstances. In other words, it tries to make the outcomes for those individuals having the same effort as equal as possible. For example, two people from different parental education backgrounds will receive the same outcome as long as they try equally hard. In contrast to the ex-ante approach, the ex-post approach refers to actually observed outcomes by the individuals. Individuals are grouped in *tranches* (as opposed to *types*) where they are defined by their relative position in terms of effort distributions across types and any inequality within these tranches allows to measure inequality in opportunity (FG, 2011, p. 625). This is in contrast to the ex-ante type grouping that computes IOp based on differences between types of people. The ex-post approach requires large scale and exhaustive datasets with information on the effort variable. Computation becomes heavy and complex when the effort variable is not observed and assumptions need to be made between effort, circumstances and outcome that allow estimating the effort variable.

Roemer and Trannoy (2015, p.284-289) refer to the ex-post as the *Fairness Gap Approach.* This is because the inequality estimate is the *fair* inequality arising from different efforts (i.e. ex-post in efforts). According to them the *Fairness Gap Approach* "measures the gap between the inequality of the actual distribution and the inequality of a counterfactual distribution in which all the effects of circumstantial variables have been removed, either by suppressing them, or by imputing to each individual a reference value of circumstances such as the average one. […]. If we had estimated a reduced form with only effort variables (something that has not been done in the literature so far), we could have (an) analog of formula […] with an estimation of the inequality of the expected outcomes across tranches when circumstances are in the residual and have been removed. Computing directly from the data the average outcome of those sharing the same effort, as done by Checchi and Peragine (2010), is a non-parametric way of doing this)".

The second principle is the **reward principle**. It tries to reward people for differences in effort with the same circumstances. A question posed is how much reward effort should permit. However, there is no obvious answer in the literature. The reward principle can be categorised into three sub-categories. First, the **liberal award** principle is the most widely accepted reward principle in the literature of distributive justice. It states that differences in outcome arising from divergent efforts should be respected and granted (and even encouraged). This principle states that no more compensation should be made than is needed to correct inequalities due to circumstances. In particular, it would generally follow that within type (same circumstance), those who expect higher effort will have greater welfare.

Second, the **utilitarian reward** principle requires no inequality aversion. The goal lies in maximising the overall outcome for people within the same type. In other words, it focuses on the sum of incomes of those that only differ in terms of effort.

Third, the **inequality-averse reward** is characterised by either the stochastic nature of outcome (after conditioning on circumstances) or by risk aversion. It rejects the utilitarian and liberal rewards, because compensation may be necessary even after taking circumstances into account and other factors might affect income (e.g. luck, differences in risk aversion).

Roemer (1998) emphasises that the concept of IOp is not to be considered as a complete theory of distributive justice for two main reasons. First, pragmatism should serve as a guideline. However, a complete theory and measure for what and who people are, what they are responsible for, etcetera does not exist. The IOp approach is ought to provide policy recommendations for societies that have

found consensus for the underlying norms of responsibility. But this also means that the choice of types and even the scope of policies are dictated by social norms. Second, the IOp theory does not provide guidelines on what proper rewards for efforts should be.

As briefly discussed in this section, the decision of the underlying principles is important for choosing an IOp methodology that corresponds and maximises the potential of the dataset at hand. Section 3.1.6 elaborates on the frameworks and main models in the IOp literature. The model in this dissertation follows the compensation principle's ex-ante approach by FG (2011).

3.1.4. Inequality of Opportunity: Challenges and Trade-Offs

While Roemer's IOp concept is appealing to many and has certainly awakened the interest for policy-makers, its operationalisation for empirical research is challenging. Moreover, Kanbur and Wagstaff (2014; 2015) argue that it may not be the most suitable analytical tool to instruct equality policies due to the lack of consensus of its underlying philosophical, ideological and normative thoughts.

A series of questions come to mind when investigating into the IOp concept. When should an aspect of the environment be placed as a circumstance and what is really out of one's control? To what extent are we responsible and accountable for our behaviour and actions? Should playing fields be levelled-off partially or fully? What is unfair and what is merited? In order to reply to these questions, it is important to isolate the various underlying elements that guide this rhetoric. Three sets of simplified trade-offs can be distinguished, yet they partially overlap.

First, a clash lies in the normative and ideological debated that should formulate the concept and consequences of IOp and the fact-based positive approach to it. Second, this then entails concerns about the theoretical appeal of the IOp concept, which stands in opposition to the degree of its operationability. With these elements in mind, third, it is important to try to draw a line between the equality of opportunity policies that should strive to be effective and efficient in boosting personal fulfilment and aggregate development, and the purpose to maximise equality in an attempt to guarantee equal opportunities for all.

I. Normative vs. positive debate

What is fair or unfair? What are the underlying normative thoughts behind inequality and inequity? Barros et al. (2009) summarise that until the 1970s the common understanding of assessing fairness and equity of social allocation was primarily based on the distribution of outcomes. Rawls (1971) was one of the first to question the until then narrow views of fairness in his book on the "Theory of Justice" introducing elements of opportunity, liberty and primary goods that allow to level playing fields in Roemer's terms. Since then much debate and literature has emerged on the Theory of Distributive Justice (e.g. Rawls, 1971; Sen, 1985; Roemer 1993; Cohen, 1989; Salverda et al, 2009; Barry, 2005).

It is important to note that a thorough analysis of the normative and ideological questions are out of the scope of this dissertation. Yet, a brief insight is given in order to understand the complexity of the IOp concept and its utility for policy makers (the next section elaborates more on this). The extent to which a normative judgement can be given (e.g. how and to what extent income should be redistributed or opportunities equalised) largely depends on the specific cultural, socio-economic,

political and jurisdictional context of the society to which the IOp analysis is applied and must be debated publicly in order to find a consensus and support.

Roemer's (1998) normative propositions for his IOp concept and EOPs are mostly acknowledged in the literature, because he refers to general concerns of human values, such as the proposition that all humans are equal and thus we should all be treated equally and that every human should have access to basic education. However, to empirically analyse his theoretical model and to then execute EOPs based on the results is difficult.

At first a consensus must be reached on which theory of justice should be used. Much of the IOp literature finds a consensus that opportunities should be to some extent equalised, that people should not be discriminated based on their environment and that an important way of doing so is by levelling-off playing fields through, for instance, readily available access to high quality education. But, how to pragmatically achieve equal opportunities, until what age and at what price education should be made accessible, and so forth, is barely discussed. This is, because it is up to each society (or political leaders) to decide what they view as the most appropriate.

Two often referred to theories of justice and welfare are the theories of Utilitarianism and Rawlsianism. In brief, utilitarianism has the goal to maximise the utility for the maximum amount of people. Rawlsianism aims at maximising the utility of the weakest. In other words, social welfare only increases if the utility of the weakest (i.e. the most disadvantaged) increases. This is commonly known as the MinMax social welfare function. So, already by looking at these two very different theories of distributive justice one can see the diverging direct implications to the corresponding nature of EOPs.

The endlessness of normative debates that are shaped by values, morale, norms and personal experiences stand in contrast to the positive debate of economists who tend to focus more on efficiency than equity and a fact based analysis. One could even argue that even if the equity debates were agreed on, many would still focus on efficiency aspect. One example would be the 2007 financial crisis. Bailing out the big banks in Europe was hugely unfair, but efficient. Many who do not want to tackle the messy theory of welfare economics tend to stick with empirical research. So, it may also be argued that first the facts need to be analysed and then the solutions to them (and the underlying) norms can be discussed. For instance, the opportunity deprived have higher returns to education, therefore it would make sense to disproportionately invest into their education.

But in the end, this is like the debate of what came first, the chicken (norms and values) or the egg (facts and figures)? Yet, the argument of first doing fact digging is easier said than done, particularly in the IOp field, where the general theory is difficult to operationalise and is based on the underlying perception of justice. In sum, it is difficult to make a clear distinction between the normative and positive debates, because they are heavily interlinked.

II. Theory vs. empirics

The IOp concept is appealing, but applying the concepts to data in an attempt to quantify efforts, opportunity deprivation and an inherent sense of discrimination can be difficult. The effort variable, for instance, can be defined in broad terms such as active decisions taken, the subjective valuation of preferences, the time spent doing homework or building one's social capital. Attributing a numerical value to these intrinsic elements is difficult. For one person, the value of spending more time with the family may be equally worth as for someone to work overtime and get a higher salary and promotion. Then the consequent differences in income should be perceived as fair and the intrinsic valuation of utility should be equal.

But how are these preferences measurable? With the difficulty of measurability of efforts comes the lack in data availability that allows such "debatable" measuring capabilities. Furthermore, to what extent are those decisions taken by an individual actually independent from the environment? To what extent can we be taken accountable and responsible for the decisions we take in life? Roemer

(1998) iterates the significance and definition of each of these two terms. Section 3.1.6.2 elaborated more on the independence assumption between circumstances and efforts. So much of the difficulty in operationalising the relatively intuitive IOp model lies in, for instance, value (in)tangibility (e.g. how to measure advantage), information asymmetry (e.g. can efforts be observed) and statistical availability. It is these operationalisation challenges that have led to a broad literature of, but yet little consensus on the empirical models of IOp.

Nevertheless, some economic sectors were quick to study concept of IOp and attempting to attribute values to the effort and responsibility variables. The health and education sectors are examples of such. Jusot et al. (2013), for instance, study the relationship between efforts and circumstances in the health sector and investigate into the extent to which smokers should be eligible for state sponsored lung transplants and the extent to which smoking is self-inflicted.

The next paragraph discusses some of the efficiency versus equity debates and the next section looks closer at the links between the analysis of IOp and public policy.

III. Efficiency vs. equity

There are important equity-efficiency trade-offs in the concept of IOp and EOPs, many of which are due to divergences in returns to investment. Roemer (1998) mockingly discusses some of these with the goal to emphasise that the normative principles of equity should not be forced without sufficient efficiency rationale. EOPs can try to target efforts (e.g. through encouragement and reward) and circumstances (e.g. by discrimination alleviation and levelling-off playing fields).

On the effort side, should a short person who tries really hard be allowed to play on a national basketball team? Or should a medicine student who tries really hard, but keeps failing the medical exams, be allowed to operate on humans?

On the circumstantial side, should all have equal access to education and all pass the high school exams? For the broad average of the population yes. But for a mentally handicapped person to reach this goal a disproportionate amount of public resources may be required, while a super smart kid with a genetically high IQ may be underchallenged, thus lacking fulfilment. This is what is referred to in the literature as the *Problem of Slavery of the Talented* and the *Social Biological Problem* (Roemer, 1998). Returns to investment are obviously quite different for the extreme ends of the spectrum questioning the efforts to maximise equity at all costs.

Roemer (1998) argues that more than half of the disadvantaged who receive opportunity-egalitarian policy do not actually improve in status, or even see their situation worsen. It is thus crucial for EOPs to choose adequate measures for success and advantage, type definitions, effort measures, sets of admissible policies, allocation rules and so forth. Thus, the next section discusses the importance and relevance of the IOp concept for the public and policy-makers.

3.1.5. Inequality of Opportunity and Public Policy

Should one worry about inequality? Should policy makers object to it? And how does the concept of IOp fit into these debates?

Much literature exists on the role of inequality in development. Particularly the study of the relations and causality of the poverty-inequality-economic growth triangle has been subject to heavy ideological, empirical and operational debates (e.g. Bourguignon et al., 2007). The inequality components of this literature however focus primarily on inequality in outcome. As subsequently explained, it can be argued that the shifts in public opinion of the "justice" of inequality should stimulate a revisit of the poverty-inequality-economic growth triangle with an IOp component and subsequently lead to an adaptation of the development policies put forward.

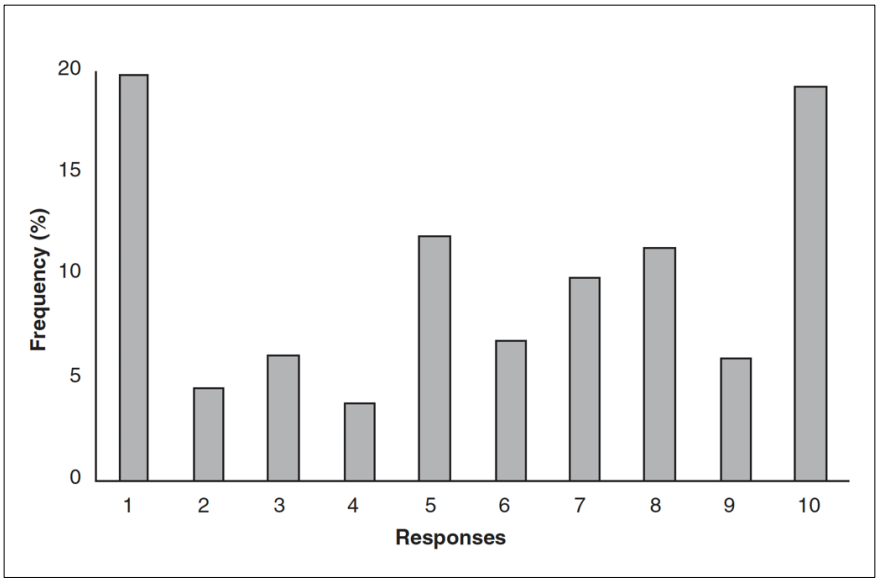
3.1.5.1. Importance of Inequality of Opportunity for the General Public

The literature on the theory of distributive justice and public opinion polls suggest that there has been a shift in ideals when referring to inequality (e.g. Salverda et al, 2009; Roemer, 1993; FG, 2011; Barros et al., 2009). Inequality in outcomes arising due to difference in efforts are increasingly seen as being more socially acceptable and fair. Thus, great efforts should be made in levelling-off playing fields by minimising the effect that exogenous socio-economic conditions have on outcome and efforts should be incentivised in order to maximise everyone's potentials and fulfilment and subsequently boost overall aggregate economic performance. This logic is increasingly present in the developed world and on the rise in emerging markets. Figure 3.4 illustrates a globally representative opinion poll cited by Barros et al. (2009, p. 26) confirming exactly that trend. Individuals in a representative population sample from over 69 countries were asked whether they would prefer more or less inequality. Possible answers ranged from 1 (equivalent to "incomes should be more equal") to 10 (equivalent to "we need a larger income difference as incentive for individual effort"). Interestingly, an almost equal share of respondents (20% each) replied in saying that incomes should be made equal and that incomes should be the lead to incentivise effort.

Politics has been fast in picking up the rhetoric of equal opportunities in an effort to win votes. One example would be the call of the German government placing the concept of "Chancengleichheit" (i.e. Equality of Opportunities" in English) as part of the core of the governments' coalition agreements since the early 2000s (Government of the Federal Republic of Germany, 2013).

FG (2011) argue that there are three principle reasons as to why equal opportunities are important for the general public. First, surveys show that individuals prefer equal opportunities to other principles. Second, the IOp concept tries to identify what depends on an individual's circumstances and what on efforts encouraging the principles of meritocracy that rewards efforts. Last, the study of EOp is important for numerous jurisdictional reasons, in debating norms and customs, but also identifying the degree of personal accountability and responsibility for taken decisions and actions.

Figure 3.4: Should We Have more or Less Inequality?
Responses From the World Values Survey



Source: From Barros et al. (2009, p. 26). Survey conducted by the Inter-University Consortium for Political and Social Research, based at the University of Michigan, 1999–2000, as cited in Inglehart and others, 2004.
Note: "1" is equivalent to "incomes should be made more equal"; "10" is equivalent to "we need larger income differences as incentives for individual effort."

3.1.5.2. Importance of Inequality of Opportunity Analysis for Policy-Makers

The previous section described a general shift in public opinions on the perception of what is considered unfair or bad and what is considered more socially acceptable and not so bad inequality. Politics has picked up on this in their electoral rhetoric in some of the developed country's democracies emphasizing the importance of equal opportunities for all.

Other sections of this chapter, such as the IOp literature review or the model description, explain why IOp can be considered as a measure of *true* inequality and therefore serve as a more adequate tool to facilitate the design of effective public policy than the conventional concept of inequality in outcome.

The World Bank (2006) argues that "income inequality due to circumstances may lead to a suboptimal accumulation of human capital and thus to lower growth". Its 2018 report on human capital entitled "The Human Capital Project" further builds on this. While not explicitly citing its previously developed Human Opportunity Index analyses, it focuses on the EOPs that are ought to eliminate the effect of socio-economic circumstances (World Bank, 2018). This clearly highlights the importance of the study of IOp and its opportunity depriving circumstances for the policy sphere.

Raj Chetty, a well-known economist and specialist on inequality studies, has initiated the successfully expanding Opportunity Insight Project, which "uses big data to empower policy makers and civic leaders to create targeted local policy solutions that revive the American Dream" (Opportunity Insight Project from https://opportunityinsights.org). In other words, the platform encourages empirical research and large-scale data analysis in order to isolate the most opportunity deprived communities and groups of people in the US and consequently contribute, shape and encourage policy-makers to employ localised and customised effective EOPs. This project is a great example of how civic engagement is actively beginning to shape opportunity enhancing policies at the local and at the aggregate level.

Jusot et al. (2013) argue that not only is the general concept of IOp appealing to the general public, but it is of particular interest to policy-makers, and especially for the health industry. The health sector, in fact, was one of the first domains eager to incorporate IOp research into its work. This is in part because of the rhetoric of unequal access (or unequal opportunities) to health care and its public goods nature. Jusot et al. (2013) elaborate on the example of the public and within industry debates in the UK whether smokers or alcoholic should have access to public budget sponsored lung or liver transplants. The question here is whether these diseases were self-inflicted (personal responsibility and accountability, therefore captured by the effort variable) or whether these are caused by exogenous factors, in which case public health care should largely cover the medical expenses associated with these diseases in case the person in questions is a smoker or an alcoholic.

Furthermore, Marrero and Rodriguez (2013) argue that one of the main reasons why the empirical literature of inequality in income on economic growth is so indecisive is because it combines elements of unequal opportunities (exogenous circumstances) and efforts (endogenous decisions and actions). They study the effects of inequality in opportunities (circumstances) and the effects of inequality of efforts on growth and find that there is a negative relationship between the inequality of opportunity and a positive one between inequalities of efforts and growth. They thus argue that the IOp research is of direct relevance for policy-makers, who should focus on discrimination elimination and effort enhancing policies.

Being able to identify the factors that impact most on individuals' opportunities makes the study of IOp relevant for policy-makers, individuals, economists, but also specialists from other academic fields, such as sociologist or anthropologists. According to FG (2011, p. 623) there are three main objectives as to why the study of IOp is crucial. First, the design of public policy must take into consideration IOp. More and more consensus exists that unequal opportunities, and not inequality in outcomes, should serve as a guideline for policies. Social and economic policies should focus on eliminating inequalities that arise from unequal opportunities and not from unequal outcomes. This is because opportunities are largely seen as out of one's influence and personal responsibility.

Second, in order for these policies to be effective, public attitudes must divert away from outcome to opportunity inequalities. There must be a redistribution of public and private resources. FG (2011) highlight the importance of attitudes and beliefs in generating multiple equilibria with different objective economic characteristics.

Last, the study of IOp helps to understand whether more equal or unequal societies demonstrate better or worse aggregate economic performance and the study of IOp can contribute to this discussion.

Wagstaff and Kanbur are two economists who in general are relatively critical of the utility of IOp for policy-makers (e.g. Kanbur and Wagstaff, 2014). In their 2015 paper however, they explain what is needed to turn the analysis of IOp into a useful reference for policy-makers. They agree that not all inequality is bad, in other words, inequality that arises due to different efforts and decisions should even be encouraged. However, bad inequality, in other words, inequality that arises due to differences in opportunities (i.e. exogenous factor discrimination) should be eliminated or at least an attempt should be made to reduce them to a maximum. So, it is put forward that the analysis of IOp should only serve as a tool for policy-makers when one can isolate the opportunity deprivation share out of overall inequality. This is exactly what the IOp indices in this study aim to do. Their interpretation and particularly the results interpretation of the inequality of opportunity ratio (IOR) is the share of opportunity deprivation of overall inequality in economic welfare. Whatever is not captured by IOR, can be considered as *legitimate* or *acceptable* inequality.

To sum up, the study of IOp is important, because public attitudes have changed increasingly acknowledging the difference between fair and unfair inequalities. Also, the study of IOp allows to isolate the most opportunity deprived groups of people and hence facilitates the formulation of target oriented EOPs, which in turn should boost aggregate (socio-)economic performance.

3.1.6. Inequality of Opportunity: Frameworks

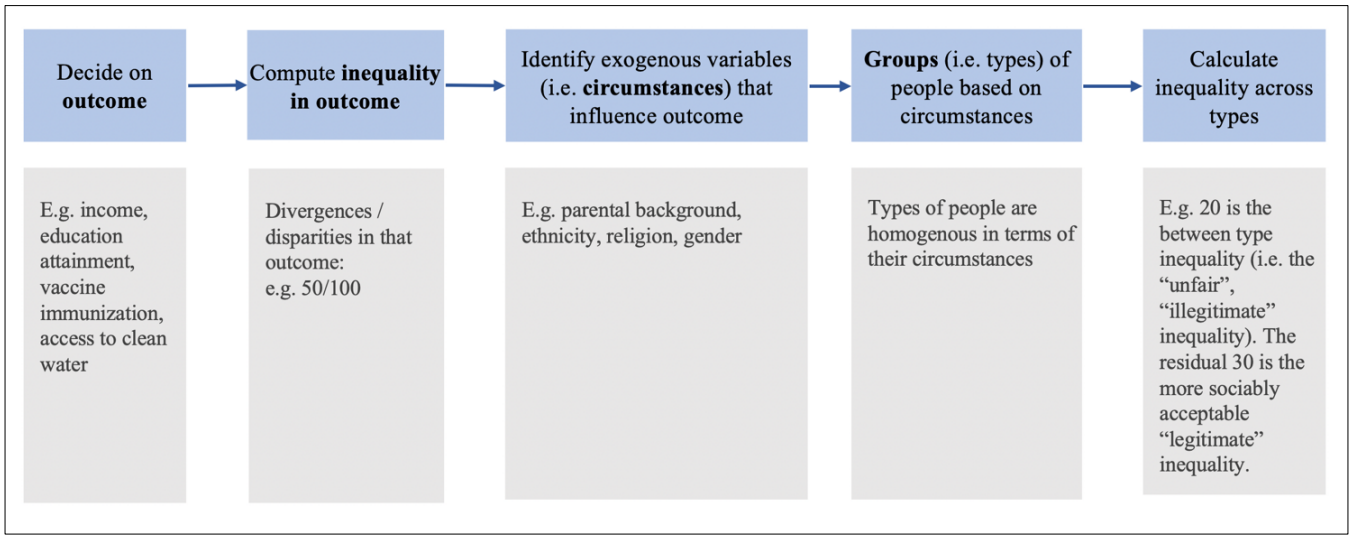
Setting aside all technicalities, Figure 3.5 is a flowchart that summarises the general methodology when studying IOp. The chart shows five principle steps to computing IOp indices all of which are somewhat relevant for all different approaches to studying IOp. First, an outcome measure needs to be decided (e.g. income, consumption, education attainment, access to the credit market or water supply). Second, inequality in outcome needs to be computed using an adequate inequality measure. For example, we could see that income inequality is 50 to 100. This dissertation has dedicated an entire chapter for justifying the choice of outcome variables (income and consumption) and choosing the most appropriate inequality in outcome measure (Mean Logarithmic Deviation). Third, the exogenous circumstance variables need to be identified. Despite all the different approaches to tackling IOp, most papers dedicate some effort in justifying the choice of circumstance variables, as does this dissertation (Section 3.3.2.5: Circumstance Variable Definitions & Adaptations). Fourth, individuals need to be grouped into types of people in which they are homogenous in terms of circumstances (or efforts). This dissertation highlights the descriptive statistics for types and type distributions for all levels of the analyses. Last, the IOp indices are calculated. For example, if

previously inequality in outcome was judged as 50 to 100 and the IOp computation indicates that 20 of the 50 in inequality in outcome is due to between type inequality (i.e. unfair inequality due to circumstance discrimination), then this would mean that the remaining 30 is due to the socially more acceptable legitimate inequality (i.e. arising from differences in effort).

Many approaches are used for estimating IOp, both at the theoretical and empirical level. The previous section on the IOp principles gave a first insight into the divergence of approaches. This section further elaborates on this from a more pragmatic view by comparing the three main strands of empirical model approaches. As already mentioned, this dissertation uses the FG (2011) direct ex-ante approach to study IOp in Nepal.

On a side note, this dissertation only covers cross-section models since the operationalisation of the few existing theoretical panel IOp models is virtually impossible. This is primarily due to the lack of data availability. Section 3.10.1 (on this study's limitations and possible extensions) provides a brief insight into a theoretical panel expansion and some empirical difficulties for it.

Figure 3.5: Simplified Intuitive Flow of IOp Methodologies (In General)



Source: Author's elaboration based on description by Wagstaff and Kanbur (2015)

3.1.6.1. Inequality of Opportunity Models: Cross Section Analysis

According to FG (2011) there are **three main approaches to measuring IOp**. The first approach is a linear regression based model that decomposes IOp in terms of effort and circumstances. The second approach gives an ex-ante and an ex-post decomposition of IOp. The last approach uses a stochastic dominance comparison of distribution of different type distributions in order to measure the degree of IOp.

Furthermore, Table 5.3 in the appendix summarises the main empirical IOp literature. It also contains information on the chosen approaches for estimating IOp. As can be seen, a majority of studies employ the approach used in this dissertation, namely the direct ex-ante approach. For these, often the study by FG (2011) serves as key reference.

Approach 1 is a linear regression based inequality decomposition (e.g. Bourguignon et al., 2007). Professor Bourguignon has long been in the centre of debate on the issue of inequality decomposition. In their paper, the linear model of outcomes is a function of effort and circumstances. The circumstances (elements of opportunity) are then further decomposed into indirect and direct effects that affect choices. The indirect effects are supposed to capture the extent to which effort is

influenced by circumstances. Furthermore, they make use of hypothetical distributions in order to be able to make assumptions on the effort variable and supressing the effect of circumstances. In their IOp analysis for Brazil, they find that their five circumstance variables (mother's and father's education, birth region, race and father's occupation) explain between 10 % and 37% of the Theil Index (their choice of inequality in outcome measure; direct and indirect effect) and 6% is due to the direct effect of circumstances on earning inequality. Parental education is found to be the most important exogenous variable in their study (though father's occupation and race are also influential determinants of unequal earnings). Since a linear regression model is applied, only a limited number of circumstance variables are used and therefore there is great likelihood for omitted variables. Furthermore, in their first regression the key (questionably too strong) assumption for them to be able to infer causality is that elements of efforts are not influenced by circumstances. Regression based inequality decomposition is highly technical and results can be interpreted in different manners. Also, their approach requires large scale and relatively exhaustive data. Yet, Bourguignon et al.'s (2007) IOp decomposition method remains the most frequently used in the field of opportunity economics. As will be explained later, the FG (2011) approach to study IOp bears a few advantages over the Bourguignon et al.'s approach. One example is the FG (2011) approach does not have to adhere to the independence assumption between circumstances and efforts, because it does not require to observe effort. It may therefore be argued that the FG (2011) approach is more solid and provides more robust results. However it does not allow to infer causality, but only correlation.

Hassine (2011) runs a similar approach to Bourguignon et al. (2007) in her study on Egypt, which she claims is the first empirical research on IOp in the Arab region due to limited HH level data availability. She evaluates the contribution of IOp to earning inequality by using parametric and non-parametric estimates of a lower bound for the degree of IOp for wage and salary workers. She does this for three time periods (1988, 1998, 2006) and different populations. She also accepts the independence assumption that circumstances have no direct causal effect on efforts (or vice versa), which allows her to run a regression based IOp decomposition and determine the reasons for differences in earnings. Without the acceptance of this independence assumption one cannot conclude on any causal relationship. But she also elaborates a model without effort, allowing her to relax the independence assumption. She uses six different circumstance variables (mother's and father's employment and education, birth region, fathers occupation status) and finds that earning inequalities due to opportunities declined from 22% (1988) to 15% (2006) and that birth region and father's background most effect earnings. She also found that the effect of mother's education has significantly increased over the studied time period.

Approach 2 (e.g. Checchi and Peragine, 2010) is a standard between-group inequality decomposition that estimates IOp through an ex-ante (circumstance) and an ex-post (effort) method. However, the principal difficulty in their study is to clearly differentiate and isolate circumstances and efforts. FG (2011, p.629) give a complete debate on this approach. Checchi and Peragine (2010) use relative income as effort variable and gender, birth region and parental education as circumstance variables. They find that particularly parental education effects unequal opportunities, which account for over 33% of income inequality for all of Italy. Gender is found to be an important factor for IOp in Southern Italy, while not as important for Northern Italy (seen as more developed than Southern Italy).

Furthermore, Asahulla and Yalonetsky (2012) utilise three different indices that measure the equality of educational opportunity in India, based on the ex-ante and the ex-post approach. They highlight that intergenerational correlations serve as imperfect indices of inequality for two main reasons. First, they relate a limited set of circumstances beyond the individuals control for economic outcome. This then overweighs welfare inequality for which individuals are not accountable. Second, when considering well-being conditioned by circumstances, distribution differences are deemed to contribute to IOp and intergenerational correlations are inappropriate to measure IOp. However, they

also accept the general approach by Roemer (1998), where IOp is divided into circumstances, effort and advantages. They utilise three different indices in order to assure for robustness, all of which give similar results and draw particular attention to father's background and geographic origins. Mother's education impact rose over the studied time period and there was an increase in IOp across age groups. The Pearson-Cramer (PC) Index was the first applied index by Asahulla and Yalonetsky (2012). It is sensitive to different group size and migration within a type and it weights the average probability of the population achieving different earnings across different types. The second index is the Overlap Index, which compares "representative agents" from each and every group, independently of seize (unlike the PC Index). It measures between-group inequality of outcome distribution. Lastly, they use two measures of the Reardon Index. The first is a cumulative probability of achieving a certain earning level conditional on belonging to a specific type. This, however, makes the results largely depended on the populations' cumulative probability. The second measure of the Reardon Index is neutral to subcategories. It is an unweighted sum of within-type inequalities. Both of the Reardon Index measures are exceptionally sensitive to the relative desirability of the level of the circumstances categorical variables such as education attainment. There is minimum inequality when all individuals have the same outcome and maximum inequality when all individuals are evenly split between worst and best outcome (Asahulla and Yalonetsky, 2012).

Approach 3 (e.g. Lefranc et al., 2008) is a *stochastic dominance comparison of distribution* conditional on types. This approach is fairly new and very complex. They also compute a *Gini of Opportunity Index* as a scalar measure of IOp. This approach is not further discussed since it goes beyond the scope of this study (which focuses on the concept of IOp as defined by Roemer) and is not used in the empirical analysis.

FG (2011) merge approaches 1 and 2. From Bourguignon et al. (2007) they utilise the parametric inequality decomposition and from Checchi and Peragine (2010) the non-parametric ex-ante method. The FG (2011) model is described in detail in Section 3.2, but a brief summary is given here.

FG (2011) only utilise circumstances and do not include effort variables in their analysis. This has the imminent advantage that the independence assumption between circumstances and efforts can be dropped. As mentioned in the section on IOp principles, FG (2011) use the ex-ante compensation principle for studying IOp. It states that EOp is said to exist when everyone has the same opportunities, despite differences in circumstances. With this postulate, effort does not need to be observed (or measured), because expected or hypothetical outcome levels are computed.

For the parametric approach, FG (2011) first estimate the linear model effects of circumstances on economic advantage (like Bourguignon et al., 2007). For the non-parametric approach, FG (2011) refer to the ex-ante approach by Checchi and Peragine (2010). To put simply, expected or hypothetical outcome levels are computed. Then the difference between the hypothetical and the actual outcome distribution inequalities are calculated. From there, in both the parametric and non-parametric approaches, two scalar IOp indices are computed. First, an absolute measure of the level of inequality of opportunity (IOL) and second, a relative measure of IOp to overall outcome inequality (IOR) are calculated. The results are interpreted as lower-bound estimates on the share accounted for by all circumstances. Both estimates are robust results for large samples, but the parametric approach leads to more moderate estimates in smaller samples. FG (2011) explain that the parametric approach is preferred, since it allows for a more specific relationship than the non-parametric approach. Also, reduced-form ordinary least square (RF-OLS) circumstance estimates are provided, since they require fewer assumptions than non-linear models and may therefore, to some extent, be more robust. Furthermore, the specific circumstantial shares are computed referred to as *partial IORs*. The RF-OLS regression results and the partial IORs also give an indication as to which exogenous factors have a

larger effect on economic advantage and therefore IOp and as to which groups of people are most opportunity deprived.

3.1.6.2. Inequality of Opportunity: Methodological and Measurement Challenges

There are several measurement issues when trying to operationalise the IOp concept. Principle causes are the lack of data availability, information asymmetry, valuation and divergence in underlying norms. Roemer and Trannoy (2015, 2016) and Ramos and Van De Gaer (2012) elaborate extensively on some of these main challenges. Emphasis here is placed on the following issues that are the most relevant for this study.

I. Multidimensional concept and reward/compensation norms
II. (Im)possibility for causal inference
III. Inclusion and measurement of effort

I. Multidimensional concept and reward/compensation norms

The concept of IOp is multidimensional and there is little agreement on the normative judgements behind it. This was previously shown in the principles section. The appeal of Roemer's basic IOp approach is the theoretical simplicity of clustering the complexity of inequality in outcome and the various causes of it into essentially two components, efforts and circumstances. Roemer and Trannoy (2015) argue that first, there is no consensus to what degree and how efforts should be rewarded. Second, there is no consensus to what degree and how circumstances should be compensated. Most would agree that the most disadvantaged should benefit of some sort of EOP schemes (e.g. through education programs), yet the most advantaged can be left as they are benefitting from their environment. Third, circumstances and efforts are heavily interconnected, which makes the evaluation of inequality difficult. So this further renders the judgement on whether and how to reward effort problematic.

II. (Im)possibility for causal inference

Two main elements continuously emerge in the literature on the difficulty to make causal inference. First, the matter of reverse causality can be largely disregarded when looking at the circumstance variables when computing the IOp estimators. This is self-explanatory because outcome is unlikely to affect the exogenous factors. Because this study excludes the effort variable, no reverse causality is in question for efforts.

Second, the existence of endogeneity due to omitted variables is an issue for all IOp approaches. This is largely due to a lack of data availability and information asymmetry. One often cited example is the situation of "luck" (e.g. see Ramos and Van De Gaer, 2012) or the "genetic" variable that allows for talent and skill. Bourguignon et al. (2007) argue that "an instrumental variable strategy is unlikely to succeed, since it is difficult to conceive of correlates of the circumstance variables that would not themselves have any direct influence on earnings."

Despite the fact that the lack of being able to make causal inference when calculating IOp, Roemer and Trannoy (2015) emphasise the importance of the correlation and relationship trends that IOp analyses allow such as being able to say what circumstances correlate with the degree of IOp and how this compares to the overall level of outcome inequality.

III. Inclusion and measurement of effort

While the broad and general definition of efforts in the IOp concept is appealing at first, the pragmatic operationalisation of it bears many challenges and has given rise to many econometric approaches trying to capture it appropriately (i.e. its measurement and its impact on inequalities analyses). Five main challenges with the effort variable are highlighted.

i. Effort has a multifaceted nature
ii. Effort information is private (lack of data)
iii. Effort is difficult to measure and quantify (information asymmetry)
iv. It is difficult to find suitable proxies for effort
v. Effort and circumstances independence assumption is too strong

What is effort? And how can it be quantified? When should an aspect of the environment be placed as a circumstance and what is really out of one's control? To what extent are we responsible and / or accountable for our behaviour and (in)actions?

i. Effort has a multifaceted nature

Effort has a multifaceted nature and authors include different elements into it such as decision-taking, effort working or studying, effort to build social capital, ambition, or the will and dedication to seize opportunities. Other terms such as *responsibility* or *preferences* are often used synonymously to effort in the IOp literature.

One can argue that the definition of effort depends on two main views of responsibility (Roemer and Trannoy, 2015). The first view includes the element of access, in other words what decisions and actions an individual can take or are available to him. These are both of physical and, indirectly through circumstances, of psychological constraint. The second view incorporates the element of responsibility, where individuals should be held responsible (and accountable) for their own decisions and actions (i.e. their preferences).

The divergence in the definitions of effort increases the difficulty to find adequate ways to measure it.

ii. Effort information is private (lack of data)

Roemer and Trannoy (2015) highlight the fact that surveys mostly do not include information on what would be considered as effort, because this information is highly personal. Instead, proxies must be used to measure efforts in an incomplete manner. More on this below under iii.

This study does not include effort, because there is not sufficient information on it in the NLSS dataset. The only possible proxy that could be used would be education attainment (i.e. years of schooling). Sub-section iv. discusses the difficulty in finding suitable proxies for effort including education attainment. Furthermore, the descriptive statistics and the section on circumstance variable definitions show, however, that education attainment in Nepal is still very low, meaning that the variance in education attainment would not be large enough to carry out a reliable IOp analysis including effort. Also, the imposed age restriction on the data sample (30 to 49 years) lowers the education attainment variance even further.

iii. Effort is difficult to measure and quantify (information asymmetry)

As previously explained, the definition of effort is multifaceted and there is no consensus in what elements should be included into its definition and as an aggregated variable. Furthermore, the lack of data availability constraints the measurement of effort. An additional challenge in the measurement when limited data is available is the subjectivity of the valuation of preferences (asymmetric information). Everyone's utility function is unique giving different optimal sets of choices, thus leading to unique trade-offs and valuations for work (e.g. career) and life (e.g. family) balance for example. Quantifying this accurately is extremely difficult. Despite the existence of econometric tools to try to capture this, it can be argued that this can induce more dispersion to the picture than necessary and it may therefore be beneficial to eliminate effort from the IOp estimation. For this the ex-ante approach used in this dissertation bears significant value-added, since it does not require the inclusion effort for its IOp estimation.

iv. It is difficult to find suitable proxies for effort

Since there are great challenges in capturing and measuring effort, there is need to find suitable proxies. Often, however, they prove to be mediocre.

Two often debated and used proxies for effort are educational attainment in years and hours worked. First, Romer and Trannoy (2015) consider that **hours worked** only serve as an "acceptable" proxy for when one is self-employed, but not for when one is a wage earner. This is because when one is employed the number of hours worked do not necessarily correspond to the "desired" amount of hours. Furthermore, much employment, particularly in developing countries such as Nepal, is seasonal or part-time. Part-time employment may be considered as involuntary, unemployment as bad luck and overtime as an imposition by the employer rather personal choice.

Furthermore, the arguments previously raised to highlight that "hours spent working" is not a suitable proxy for effort is further amplified in a developing country context where the share of the population that works in regular income work is very low. This dissertation does therefore not look into utilising hours worked as a proxy for effort.

Also, the "effort" or time spent doing work does not necessarily reflect the (output) productivity of the invested time. A parameter incorporating the "intensity" of the effort would need to be included, which again would be subject to great measurement difficulty and could possibly further distort the image of the IOp level.

Second, **education attainment** as a proxy for effort is almost equally debated. One aspect is that the choice of going to primary and secondary school is normally out of the child's consent. It is the parents and the law that decide for the child to go to school. So it may be argued that going to school and completing the minimum legal schooling years cannot be considered as effort. Furthermore, when a child is not actively engaged in class (i.e. being lazy), then that might also not be entirely his own independent choice. It may however be argued (Roemer and Trannoy, 2015) that tertiary and higher education could serve as a suitable education effort proxy. Yet, the choice of higher education is highly correlated and endogenous with parental education (see point v.). Also, the rate of acceptance into higher education institutions largely depends on the grades obtained in primary and secondary school, which then entails the same issues as elaborated before.

Furthermore, "education attainment in years" figures in the dataset (for all individuals), but the analysed data sample is subject to data constraints, including an age constraint (30 to 49 years). The education attainment statistics for this age cohort are arguably not sufficient for it to act as a suitable effort proxy. 1) Because only 29% of all people in that age cohort have completed education, thus leading to a very small population sample for a survey dataset that is already quite small. 2) Over 39%

of the age cohort are currently attending school or undergoing some sort of training program. Heavy assumptions would need to be accepted if these were to be included into the "finished" education attainment category and then serve as effort variable.

For all these reasons education attainment is arguably not a suitable effort proxy in this IOp study.

Consequently, both of these relatively common effort proxies are prone to measurement error and bias. Another possible proxy that is sometimes used is "hours spent doing homework". This however may also not be an entirely independent choice by the child and anyhow statistics for this do not figure in the NLSS III dataset.

Nonetheless, any included effort variable into an IOp analysis is certainly less reliable and robust than the circumstance variables.

v. Effort and circumstances independence assumption is too strong

When should an aspect of the environment be placed as a circumstance and what is really out of one's control? To what extent are we really responsible and / or accountable for our behaviour, actions and preferences?

Limited literature exists to test for the correlation and dependence between efforts and circumstances. Jusot et al. (2013) is one of the few papers that tests for the correlation between circumstances and efforts and how this impacts on the IOp (in health) estimates. Their results on the correlation between efforts and circumstances are inconclusive and they argue that due to the nature of their very specific health care IOp analysis, the correlation between the two variables is less important and the circumstance share for opportunity deprivation remains high even after attempting to control for endogeneity.

In theory, the decisions we take (on e.g. effort, choice preferences) should be independent from circumstances under the independence assumption by Roemer (1998), which means that we are fully responsible and accountable for them. For example, should one decide to place more emphasis on family life and thus climb the career ladder more slowly and therefore earn less compared to someone who prioritised the career over private life, then the consequent inequality in outcome between these two individuals is legitimate. The family person is responsible for his decision and should not complain about any income difference to the other person.

Cohen (1989) elaborates on the philosophical debate of egalitarian justice and by doing so he questions the independence between the decisions we (think we) take and the environment we live in. The fact that our circumstances and efforts are not utterly independent from each other becomes particularly apparent when we look at children. The younger they are, the more apparent this becomes. The argument is that parents' choices and efforts influence their children's actions, the way they think, the way they behave and the way they perceive themselves. Children's vulnerability to be influenced by their environment has arguably been a biogenetic evolutional success story for the human race, but it has also been the source for many civil war military groups to abuse this and engage actively in recruiting child soldiers.

If children are easy to influence, then to what extent are adults influenced directly and indirectly (i.e. subconsciously)? An area of psychological research is the sphere of self-conceptualisation. The question that arises is to what extent does our environment impact on our perception (i.e. self-conceptualisation) of our own potential? Translated into the IOp context, the question is to what extent the environment of the poorest, for instance, predicts their ambitions or motivations to for example study and work hard, strive to become a doctor rather than a pharmacists, or change the world. For long a majority of the economic development literature argued that the poor are irrational by spending irrational amounts of scarce resources on religious festivals for example. However, recent works in

behavioural economics have shown that these expenses are often rational, because they help foster socio-economic safety nets and can be regarded as an informal way of insurance (e.g. Ockenfels, 2016). These safety nets also guarantee that individuals who are part of the same group of people will always be preferred over people from outside of the group (i.e. discrimination), even if they are a more suitable job candidate, for instance. In turn, the environment then implicitly and maybe subconsciously influences the self-conceptualisation of one's own potential.

Furthermore, the research extension section of this dissertation briefly mentions the author's family tree. It highlights that despite the economic freedom to choose for over 18 generations, heir after heir chose to peruse the same professions as their ancestors. Was this intergenerational occupational immobility a circumstantial constraint or a personal choice given that most of them had complete economic freedom to do whatever they wanted?

As previously described, Bourguignon et al. (2007) run a linear regression based inequality decomposition to study the existence of IOp. The linear model of outcomes is a function of effort and circumstances. The circumstances (elements of opportunity) are then further decomposed into indirect and direct effects that affect choices. They employ hypothetical distributions in order to be able to make assumptions on the effort variable and supress the effect of circumstances. The key assumption here is the complete independence between circumstances and efforts. This has to, however, be accepted if one wants to be able to infer causality. Hassine (2011), for instance, strongly question this, so she developed both an approach to include and exclude efforts.

The ex-ante approach applied in **this study** focuses on the prospects or expected levels of economic outcome. The approach postulates that EOp is said to exist when everyone has the same opportunities, despite differences in circumstances. For example, despite belonging to an ethnic minority or majority, everyone should get the same expected outcome (i.e. same access to the same level of income). One major advantage of this approach is that efforts do not need to be observed (or measured). This is because expected or hypothetical outcome levels are computed using a) the maximum inter-types inequality approach (Elbers et al., 2008) for the non-parametric and b) a simulated regression based outcome approach for the parametric approach. More on the precise direct ex-ante model (under the compensation principle) used by FG (2011) in Section 3.1.6.

The FG (2011) approach applied in this study does not need to adhere to the independence assumption since the effort variable is not included. It may therefore be argued that this approach is to some extent more solid, especially when utilising a relatively small dataset like the one in this study. But this also means that the results of this study do not allow to infer causality, but they do give some indication on correlations and trends.

3.1.6.3. Inequality of Opportunity: Value-Added and Shortcomings of the Direct Ex-Ante Approach

There are several **reasons why this study utilises the FG (2011) method**. First of all, their indices and IOp decomposition is a relatively simple scalar measure of "true" inequality. Second, one of the main advantages is the exclusions of endogenous choice variable, which allows dropping the independence assumption between circumstances and efforts. Last, is the cunning combination of the general approaches 1 and 2, which has thus often served as a key reference for subsequent IOp research.

The FG (2011) paper has however the following main **flaws**. First, their IOp indices must be interpreted as lower bound estimates on the set of possible true ex-ante IOp. This is because it is a necessary condition for the path-independence assumption to hold, that is the direct non-parametric

ex-ante approach from Checchi and Peragine (2010) yield reliable and identical results to the parametric approach from Bourguignon et al. (2007). In other words, it is necessary "to treat the share of the inequality associated with the circumstance one observes as a lower-bound on the share accounted for by *all circumstances*, observed and unobserved, rather than as the share corresponding to those *specific* observed circumstances" (see FG, 2011, p. 625 for a detailed debate on this).

Second, one has to be cautious when interpreting the circumstantial results. The more exogenous variables are omitted, the larger the underestimation of IOp.

Third, their approach does not allow to infer causality, it merely shows the correlation and the degree to which circumstances affect output and it indicates the share of opportunity deprivation out of overall outcome inequality.

Fourth, the exclusion of effort in the IOp estimation is both of value-added and a shortcoming.

Dropping the endogenous effort variable is a key element in the ex-ante methodology used by FG (2011) and this dissertation. Doing so has numerous advantages and disadvantages. Some of the challenges with the observation and measurement of effort were discussed before. Anyhow, this dissertation is mainly an empirical extension of the IOp literature to Nepal. The heavily praised direct ex-ante model by FG (2011) is applied and adapted to the NLSS III dataset and the country specific context. It seems out of the scope of this dissertation to engage into the theoretical debate of how to observe, measure and analyse effort.

Moreover, if the inequality decomposition satisfies the path independence property, then the non-parametric direct and indirect (i.e. ex-ante (circumstances) and ex-post (efforts)) approaches provide the same and reliable results (Ramos and Van De Gaer, 2012). This is a crucial condition for the ex-ante approach from the compensation principle elaborated in Section 3.1.3 (Inequality of Opportunity: Other Principles). In other words, the same values are obtained using the ex-ante approach without effort, as would be with effort in the ex-post approach. Thus, it can be argued that efforts need not necessarily be observed.

Furthermore, by only using circumstances as independent variables however, this study takes an extreme stance by saying that individuals are not responsible or influenced by what their parents have done for examples. In other words, whether a father has worked hard in order to provide good education for the individual or that the individual has inherited the parental genes has nothing to do with the individual himself. These effort elements are captured by the residual in the analyses, which is potentially controversial and care must be given when interpreting the results. However, eliminating the independence assumption allows for relative result comparison between this and other ex-ante approach IOp studies.

Also, in the circumstantial RF-OLS regressions, only the effect of circumstances on economic outcome are analysed. One may argue that a significant share of the omitted variables is efforts which is captured by the residual and partially reflected by the low R-squared numbers (see empirical result sections). Note that when compared to other IOp studies, the R-squared values in this study are similar and sometimes even significantly higher.

Besides, the ex-ante approach also allows to investigate the presence of IOp and the degree of it in countries where datasets and their exhaustiveness are scarce. The empirical application and geographical extent of the concept of IOp is fairly sparse compared to research on standard inequality decomposition or outcome inequality for example. Its geographical coverage mainly focuses on Latin America and developed countries (e.g. Germany or US). Hassine (2011) has so far provided the only study for the Arab region. In Asia, most (of the few) IOp studies focus on India, particularly due to its caste society system, fast economic development and its relatively extensive data availability (e.g. Asahulla and Yalonetsky, 2012). The availability of extensive HH data in emerging countries is limited and the need for exhaustive data on exogenous circumstance variables and the responsibility variable poses a particularly strong barrier for empirical analysis.

In general, however, the FG(2011) paper has significantly contributed to the study of IOp on both a theoretical and empirical level and has received much praise also by Romer and Trannoy (2015). Slackening the independence assumption is favourable bearing numerous analytical and empirical advantages, despite not being able to infer causality from the results. This dissertation's analyses of IOp takes full advantages of these value-added elements and is therefore able to provide a thorough insight into the reality of social justice and economic equity in Nepal.

3.1.7. Literature Review Summary

The IOp literature review first complemented the inequality in outcome overview already given in Chapter 2. It was shown that other concepts like intergenerational mobility or meritocracy appear similar, but are distinct from IOp.

Consequently, the IOp concept, as first developed by Roemer (1993a, 1998) was described. It was his theoretical conceptualisation of IOp that gave way to a wave of rapidly multiplying theoretical and empirical literature on the subject. The appeal of his IOp approach being that factors impacting on inequality in outcome are divided into socio-economic exogenous factors (e.g. parental education, race) called circumstances and an endogenous effort variable (or also referred to responsibility or preference) for which individuals can be held responsible and accountable. He postulates that inequalities in outcome arising from differences in effort and not circumstances are arguably more just and ethically acceptable. It is this equality of opportunity principle that has caused his concept to gain wide support amongst scholars, the public and policy-makers.

The literature review then described some of the various IOp principles that have emerged somewhat uncoordinatedly since Roemer's benchmark book in 1998. The direct ex-ante approach under the compensation principle was isolated as the most suitable approach for this study in part due to the possibility of dropping the effort variable and being able to empirically apply the model to a relatively small dataset. The approach used in this study is largely inspired by FG (2011), which is fully developed in the next section.

The literature review then continued describing and explaining some of the key challenges and trade-offs with the basic IOp concept. It first looked into the normative versus positive, second into the theoretical versus empirical, and third into the efficiency versus equity trade-offs. Much of these debates are outside the scope of this dissertation. Yet a brief overview was given, because it is important to have a brief understanding of the underlying debates that are discussed in other academic fields like sociology, the theory of distributive justice or even psychology.

The following section elaborated on the importance of the study of IOp for the public and policy-makers. Indeed, there appears to be a shift in ideals in the general public to support the concept of IOp and how opportunity deprivation based on circumstances should be eliminated while efforts should be incentivised (e.g. through income differentials). Furthermore the study of inequalities in opportunities can be a point of reference for policy makers. The World Bank has widely acknowledged that "income inequality due to circumstances may lead to suboptimal accumulation of human capital and thus to lower growth". The study of IOp can contribute to alleviate this by being able to isolate the most opportunity deprived groups and enabling the design of localised and customised EOPs in order to boost aggregate (socio-)economic performance.

Last, the main IOp frameworks were reviewed. The three main strands were briefly described, highlighting the elements that this dissertation utilised (a complete elaboration of the model used follows in the next section). Then, three of the main methodological and measurement challenges in the literature that apply most to this study were investigated. These are, a) the multidimensional concept of IOp, b) the (im)possibility for causal inference, and c) the inclusion and measurement of effort. The inclusion and measurement of effort is the most substantial part that has been source of

much heated theoretical and empirical debate in the literature. It is thus subdivided into five sub-sections. i) Effort has a multifaceted nature, ii) effort information is private (lack of data), iii) effort is difficult to measure and quantify (information asymmetry), iv) it is difficult to find suitable proxies for effort, and v) the effort and circumstances independence assumption is too strong. Overall, all of these five reasons are an argument as to why the analysis in this dissertation excludes the effort variable and thus uses the direct ex-ante approach to studying IOp in Nepal, one of the main value-added of this approach.

3.2. Methodology followed for the Nepalese Inequality of Opportunity Analyses

This section focuses on the empirical IOp model used in this study and its adaptations to allow its operationalisation for the Nepalese NLSS III dataset. To do so, first the research questions and hypotheses are stated. Second, an intuitive explanation of the model and the analysis methodology is given. Third, the econometric model for calculating the IOp scalar indices and singling out the subsequent circumstantial significance on economic outcome is stated. Last, this section highlights the complementarity between the parametric and non-parametric approaches used to compute the IOp indices.

The section after this gives an insight into the country specific context, the analysed Nepalese NLSS III dataset and its shortcomings, the choice of independent and dependent variables and the econometric model adaptations.

3.2.1. Testing for the Presence of Inequality of Opportunity in Nepal: Research Questions and Hypotheses

Broadly speaking, there are four main research questions concerning the study of IOp in Nepal. The goal is to answer them as precisely as possible and close the corresponding gaps in the literature:

Q1: Does IOp exist in Nepal?
Q2: If IOp does exist, what is its amplitude?
Q3: How is IOp distributed / concentrated across the country?
Q4: Who are the groups of people who are the most discriminated against?

In an attempt to reply to the research questions as precisely as possible, the IOp chapter has four different levels of analyses, each containing rigorous descriptive and econometric elements. The four levels are:

1) National level analyses
2) Development Region level analyses
3) Urban rural area level analyses
4) Population grouping by economic welfare quarter analyses

There are two main clusters of research questions for each of these levels.

1) National level analyses

Q1: Does IOp exist? And if so, what is its amplitude?
Q2: Which groups of people are the most opportunity deprived at the national level? What are their relative mean economic outcomes?

2) Development Region level analyses

Q1: Does IOp exist? And if so, are there any differences in amplitudes between the five Development Regions of Nepal? Which region has the highest degree of IOp and which has the lowest?
Q2: Which groups of people within each Development Region are the most opportunity deprived? What are their relative mean economic outcomes?

3) Urban rural area level analyses

Q1: Does IOp exist? And if so, is there any difference in the amplitudes between the urban and rural regions? Which area faces the highest degree of IOp?
Q2: Which groups of people within the urban and rural areas are the most opportunity deprived? What are their relative mean economic outcomes?

4) Population grouping by income and consumption quarter analyses

Q1: Does IOp exist? And if so, are there any difference in amplitudes between the earner and consumer groups? Which quarter of the population faces the highest degree of within-group IOp and which the lowest?
Q2: Which groups of people within each economic outcome quarter are the most opportunity deprived? What are their relative mean economic outcomes?

This study examines whether there is IOp in Nepal by testing the null hypothesis that the scalar indices of inequality of opportunity (Inequality of Opportunity Level : IOL, and Inequality of Opportunity Ratio : IOR) are equal to zero (H_0: IOL = 0 and H_0: IOR = 0). Under the null, there is perfect equality of opportunity and circumstances do not affect economic outcome (i.e. per capita income or per capita consumption expenditure).

Hypotheses to test for Q1s above i.e. "Does IOp exist? And if so, is there any difference in the amplitude between […]":

H_0: IOL = 0 and **H_0: IOR = 0**, when IOL or IOR=0, then there is perfect equality of opportunity.
H_1: IOL > 0 and **H_1: IOR > 0**, when IOL or IOR>0, then the null is rejected and IOp exists.

In case the null hypothesis is rejected (IOL or IOR > 0), the question is to what extent each circumstance variable (C) affects economic outcome. This is done by running the following Reduced-Form Ordinary Least Square (RF-OLS) regression:

$$\mathbf{y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + u_i}$$

Where

y_i	Per capita consumption expenditure or per capita income of the individual i.		
β_0	Constant	X_{5i}	Father primary incomplete or above
β_j	Circumstance variable parameter (j = 1, … 9)	X_{6i}	Mother primary incomplete or above
X_{1i}	Female	X_{7i}	Birth region rural
X_{2i}	Ethnic minority/dominated caste group	X_{8i}	Birth region Hill
X_{3i}	Buddhist	X_{9i}	Birth region Terai
X_{4i}	Religion other	u_i	Error term

B_j represents the parameter for each exogenous categorical circumstance variable X_j ,and *i* indices the individuals under study:

H_0: $\beta_j = 0$, Circumstance *j* does not impact on economic outcome.
H_1: $\beta_j \neq 0$, Circumstance *j* impacts on economic outcome and the null is rejected.

The goal of this study is to identify which types or groups *k* of individuals are the most opportunity-deprived, based on the estimated regression coefficients of the circumstance variables that are statistically significantly different from zero ($\beta_j \neq 0$).

The above allows for the ranking of *opportunity profiles*, which represent an ordered set of types that are ranked by their mean levels of economic outcome. The opportunity-deprivation profile represents a subset of ordered partition that includes only a certain fraction of the population that belongs to the lowest ranked types.

Also, the above allows to estimate *partial IORs*, that is the circumstantial specific share of overall opportunity deprivation. Testing for the partial IORs, that is the circumstantial specific share of overall outcome inequality (i.e. IOR).

H_0: partial $IOR_j = 0$, when **partial $IOR_j = 0$,** then the circumstance specific share of circumstance *j* of overall outcome inequality is null.

H_1: partial $IOR_j \neq 0$, when **partial IOR ≠ 0**, then the null is rejected and the circumstance specific share of circumstance *j* of overall inequality in outcome is different to zero, i.e. the circumstance explains a certain share of overall inequality in outcome, suggesting the importance of the circumstantial categorical variable for the accessibility in opportunities.

The RF-OLS regression and the partial IORs add additional depth of analysis to the computed scalar indices and allow to identify social groups at which equality of opportunity policies should be aimed. Roemer (1998) refers to such policies as *Equality of Opportunity Policies* (EOPs) as a set of allocation rules that aim to maximise economic outcome for the worst-off type(s).

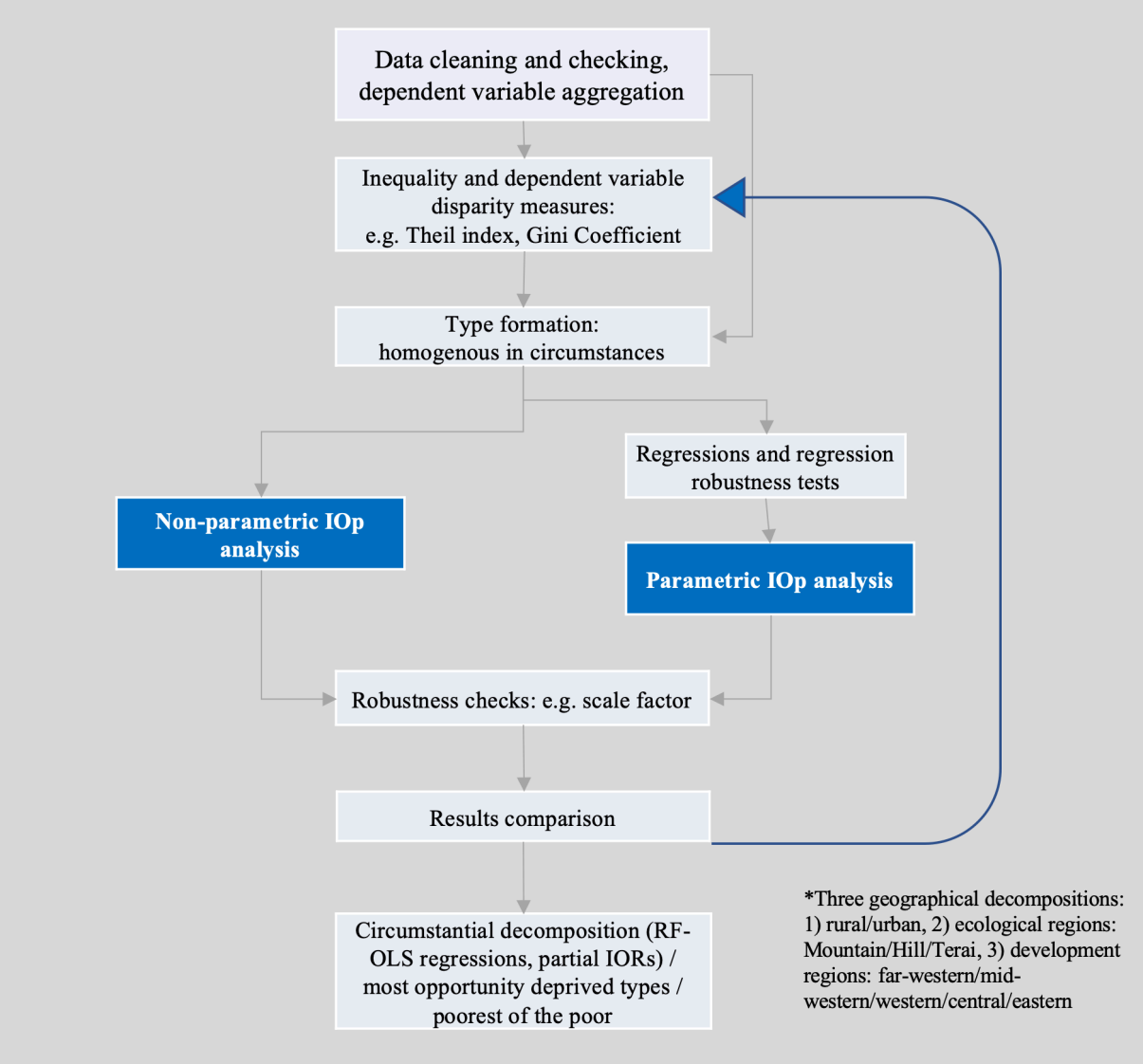
Note that the above is valid for all four levels of analyses, only that the urban rural level excludes the birth region 1 (urban rural area) categorical circumstance variables.

3.2.2. Intuitive Explanation of the Econometric Methodology

Figure 3.6 below illustrate the basic flowchart of the IOp methodology used in this study. It complements the previously discussed Figure 3.5 (Simplified Intuitive Flow of IOp Methodologies (In General)) by combining the methodological inputs taken from various authors (e.g. Roemer, 1998;

Bourguignon et al., 2007; Ramos and Van de Gaer, 2012; Checci and Peragine, 2010) and integrated into the FG (2011) model used as a main reference in this study.

Figure 3.6: Flowchart for Nepalese IOp Analyses: National and Geographical Decomposition



Source: Author's elaboration

The first two steps in the figure refer to the efforts elaborated in Chapter 2, which deal with the choice and aggregation of the outcome variable(s) and the computation of the corresponding inequality in outcome measures. This study uses per capita consumption expenditure and per capita income as the outcome variables. The chosen inequality in outcome measure is the Mean Logarithmic Deviation (MLD). Justifications for these are given in Chapter 2.

The next step calls for the identification of the exogenous circumstance variables in the dataset that are ought to (or not) impact on outcome. Once identified and categorised, groups or types of people are formed where individuals are homogenous in terms of their circumstances.

Next, two parallel approaches are taken. First, a non-parametric and second a parametric approach to computing the IOp scalar indices. A discussion on the pros and cons of each follows in Section 3.2.4 after the econometric model is formally stated. In brief here, this study uses the direct

ex-ante approach elaborated by FG (2011) and the parametric and non-parametric approaches to estimating IOp are seen as complementary.

Roemer and Trannoy (2015, p.284-289) refer to the ex-ante approach used by FG (2011) as the *Direct Unfairness Approach* (DU). This is because the inequality estimate is the *unfair* inequality arising from different circumstantial treatment (i.e. ex-ante in circumstances). They state that "[the Direct Unfairness or ex-ante approach] is computed as the inequality of the counterfactual distribution when one has removed the effect of effort variables, either by suppressing them, or by imputing to each individual a reference value of effort such as the average value. […] for the reduced form (and consequent parametric approach), a natural choice for DU is to compute the inequality of the conditional expectation of outcomes across types […] which is a neat solution chosen by Ferreira and Gignoux (2011)."

In sum, the direct ex-ante approach focuses on the prospects of outcome. Perfect equality of opportunity is said to exist when everyone has the same opportunities, despite differences in circumstances. For example, despite belonging to an ethnic minority or majority, everyone should get the same expected outcome (i.e. same access to the same level of income). One major advantage of this approach is that efforts do not need to be observed (or measured), because instead expected outcome levels are computed. In other words, the ex-ante approach measures the inequality of the average outcome across types. To put it simply, in the non-parametric approach, for instance, hypothetical standardised outcomes are computed. Then the difference between the hypothetical and the actual outcome distributions are calculated, which gives the overall Inequality of Opportunity Level (IOL). When further divided by the MLD, it gives the non-parametric Inequality of Opportunity Ratio (IOR) value. The IOR can therefore be interpreted as the level (or share) of opportunity deprivation over overall inequality in outcome.

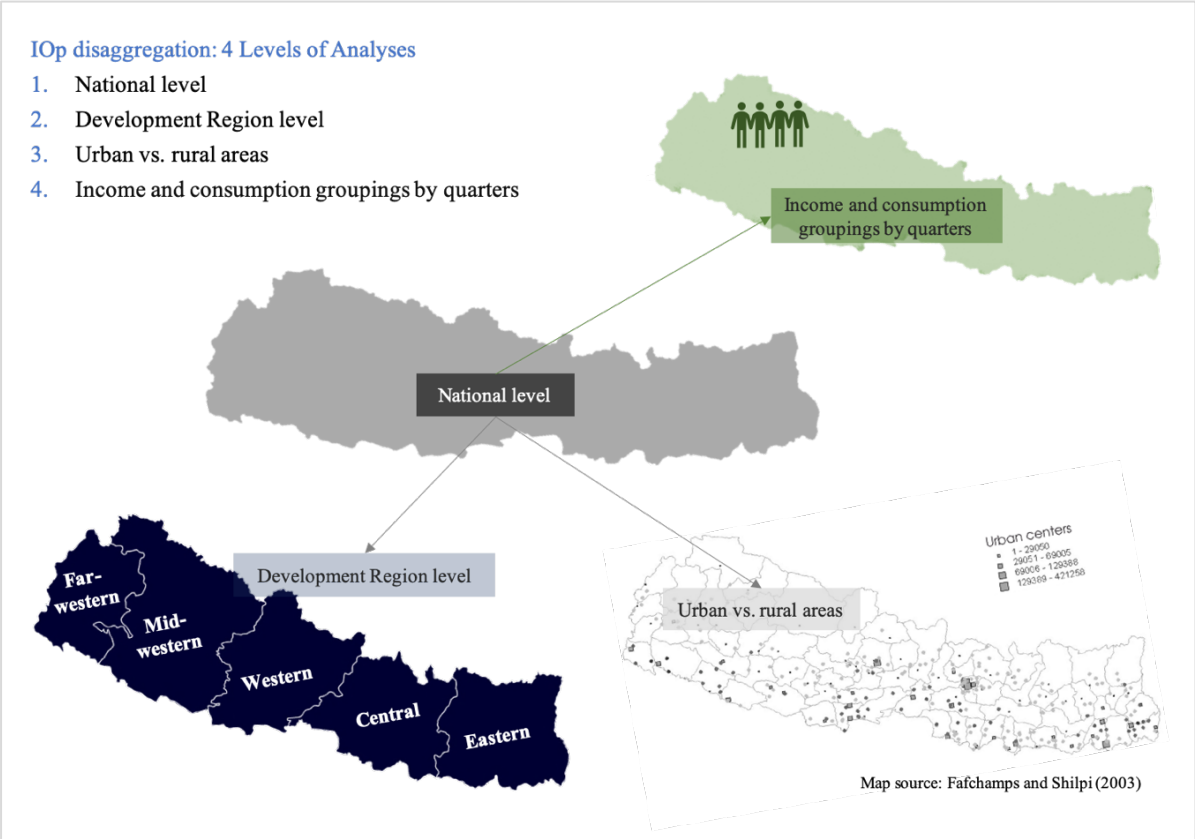
The empirical literature summary sheet (Appendix 3: Table 5.3) also shows that a significant number of empirical IOp literature employs the ex-ante approach. However, only a limited number of studies use both a parametric and non-parametric approach for estimating the scalar IOp indices. Using both of these approaches in order to check for the results' robustness is of great value added in this study. Furthermore, adhering to the relatively commonly favoured ex-ante methodology in the literature is an argument for its validity and suitability in estimating IOp.

Then, the robustness of the IOp indices is tested, primarily using equivalence scale measures (e.g. Buhmann et al., 1988; Coulter et al. 1992; FG, 2011). Note that at each stage of the analyses a maximum effort is undertaken to check the data, the process, and the analyses for consistency, plausibility and robustness (e.g. outlier checks, regression validity tests).

Next, the IOp indices results are compared to the inequality in outcome results. Then, the decomposition in circumstances follows using the RF-OLS regression model in order to investigate into the statistical significance of the exogenous factors on outcome. Also, circumstance specific shares of overall opportunity are computed, referred to as *partial IORs*. This then further allows to isolate opportunity deprivation profiles where types are ranked by their actual mean outcome, representing a ratio of the population of the lowest rank. This allows to identify social groups at which equality of opportunity policies should be aimed at. Again, these results (descriptive and econometric) are compared with the IOp indices and the inequality in outcome results.

Note that this basic flow of methodology applies to all four levels of the IOp investigation in this study. Figure 3.7 shows the four levels of analyses: first the national level analyses, second the Development Region level analyses, third the urban rural area level analyses, and last the economic outcome grouping by population quarters.

Figure 3.7: Four Levels of IOp Analyses in Nepal



Source: Author's elaboration

3.2.3. Econometric Model: Calculating the Inequality of Opportunity Scalar Indices

As already briefly elaborated on in the IOp literature review, the theoretical and empirical model used in this study is largely inspired by the approach developed by FG (2011, p. 626 - 637). Having already discussed the trade-offs of this approach, this section focuses on the econometric model.

The theoretical approach developed by FG (2011) uses Roemer's (1998) concept as its base, where advantage (or outcome) depends on exogenous circumstances (i.e. elements out of one's influence) and endogenous efforts (i.e. aspects for which one can be held responsible and accountable for). However, as discussed in the previous section, the model excludes the effort variable due to a series of reasons (e.g. difficulty in quantification and valuation, independence assumption between circumstances and efforts).

Individual i is characterised by (y_i, C_i^j, e_i). Where y is the economic outcome variable. This study utilises per capita consumption expenditure and per capita income as dependents variable. Two separate panels of analyses are run for each of the dependent variables. C represents the circumstance variables, j is the J elements of circumstance variables and e the vector of effort variables.

Individuals are divided into **population sub-groups (types)** in which they are homogenous in terms of circumstances. Where μ is the mean outcome level for k type, the key criterion is:

$$\mu^k(y) = \mu^l(y), \quad for \ k \neq 1. \qquad (1)$$

In other words, if the mean economic outcome is the same for different types, there is perfect equality of opportunity. The approach relies on the means of types rather than the entire within type

distributions. There are two arguments for this. First, much depends on the size of the dataset and type distribution and observation. Preferably, one would try to run type specific distribution functions, however, to do so a large amount of observations are needed. Unfortunately, given a) the relatively small Nepalese dataset, and b) the relatively large number of circumstance variables and therefore relatively low type distributions (see descriptive statistics sections in the IOp analyses), means that a more robust and reliable approach to estimating the IOp indices is by using type means. While both versions were tested, only the mean type distribution approach delivers robust results. FG (2011, p. 692) recommend using the approach utilised in this dissertation for small datasets of when type distributions are relatively low.

Second, the nature of the ex-ante approach used here encourages the use of type mean outcomes rather than the entire within type outcome distribution. The ex-ante approach compares inequalities arising between (rather than within) types, who share the same circumstances. This means that a) efforts do not need to be observed and b) that perfect equality exists when the mean outcome levels between types are identical as stated by (1).

One then needs to measure the extent to which $\mu^k(y) \neq \mu^l(y)$. The inequality indices are computed using a **smoothened distribution** of outcomes μ_i^k following the methodology inspired by Shorrock (1980). This means that every observation within one type receives the mean economic outcome of that type, replacing each individual advantage y_i^k with the group specific mean $\mu^k(y)$. In other words, within type advantage inequality is eliminated. From there on, two scalar indices can be estimated based on a joint distribution between circumstances and economic outcome. The two scalar indices are as follows (smoothened and non-parametric):

Absolute Level of IOp (IOL): $$\theta_a = \mathrm{I}\left(\left\{\mu_i^k\right\}\right) \tag{2}$$

IOp Ratio (IOR): $$\theta_r = \frac{\mathrm{I}\left(\left\{\mu_i^k\right\}\right)}{\mathrm{I}(y)} \tag{3}$$

The absolute or total level of inequality (i.e. total inequality in outcome) is denoted by $\mathrm{I}(y)$. As a reminder, the total inequality is computed using the Mean Logarithmic Deviation (MLD) from the Generalised Entropy measures. As described in Chapter 2, the MLD measures the percentage difference between the economic outcome of a randomly selected individual and that of the total sample average (Haughton, 2009, Ch.6).

The **MLD has two major advantages** for the computation of the IOp indices important to restate here. First, it follows the property of path independence when computing an arithmetic mean (e.g. Roemer and Trannoy, 2015; Ramos and Van De Gaer, 2012; Shorrocks and Wan, 2005; Foster and Shneyerov, 2000). Second, it is also the only measure of the Generalised Entropy (GE) indices that is decomposable between income and consumption sources and between individuals, groups and sub-groups.

First, the MLD is the only inequality in outcome index of the GE that follows the path independence property. Foster and Shneyerov (2000) give a comprehensive theoretical justification behind this and discuss the advantages and disadvantages of it for measuring inequality. They refer to the term as *path independent decomposability.* If the inequality decomposition satisfies the path independence property then the non-parametric direct and indirect (i.e. ex-ante (circumstances) and ex-post (efforts)) approaches provide the same and reliable results (Ramos and Van De Gaer, 2012).

This is a necessary condition for the ex-ante approach from the compensation principle elaborated in Section 3.1.3 (Inequality of Opportunity: Other Principles) to hold.

Roemer and Trannoy (2015) explain that the ex-ante approach measures the inequality of the average outcome across types. And the ex-post approach is "obtained by rescaling the distribution of the outcome due to effort by the ratio of average income to average income per type". This is one of many options to control for the effect of circumstances on outcome. So, if it is possible to nullify the effect of efforts on outcome, then the ex-ante and the ex-post approaches will deliver the same results, under the path independence assumption. Thus, it can be argued that efforts need not necessarily be observed.

Second, being able to decompose the MLD between outcome sources and between individuals, groups and sub-groups. It is the latter that is of particular interest for the IOp study, which divides the population into "types" or groups of individuals in which they are homogenous in terms of circumstances. The within type outcome distribution is smoothened and they are compared to the actual variance of outcome distribution and the expected level of outcome (following the methodology by Elbers et al., 2008). This difference allows for the model to neglect the non-existence of the effort variable and thus provide reliable results, especially for the non-parametric approach (Ramos and Van De Gaer, 2012).

In (2) and (3), IOR measures the level of IOL in relation to total inequality, where $\mathrm{I}(\,)$ indicates all inequality indices that satisfy the **standard axiomatic properties** for the measure of relative inequality. These properties call on the importance of symmetry (or anonymity), transfer, scale invariance, population replication and additive decomposition. If all of them are satisfied one can check that θ_a (i.e. IOL) and θ_r (i.e. IOR) satisfy all index axioms. Following are the six key standard axiomatic properties for relative inequality indices. A complete debate of these can be found in FG (2011, p. 630 – 631) and Cowell (1995).

According to FG(2011, p.631) both θ_a (i.e. IOL) and θ_r (i.e. IOR) satisfy the following properties:

i. Principle *of population*: the index is invariant to a replication of the population $\{1, \ldots, N\}$.
ii. *Scale invariance*: the index is invariant to the multiplication of all advantages by a positive scalar.
iii. *Normalization*: if the smoothed distribution $\mu_i\ \{k\}$ is degenerate, so that there is equality of opportunity, then the index takes a value of zero.
iv. *Within-type symmetry*: the index is invariant to any permutation of two individuals within a type.

Furthermore, the IOL (θ_a) satisfies:

v. *Within-type transfer insensitivity:* the index is invariant to any mean preserving spread in advantages within a type.
vi. *Between-type transfer principle*: the index weakly rises with any transfer from any individual i to j.

Moreover, FG (2011) state that because their indices satisfy the above properties and the transfer axiom (Pigou-Dalton) and because they utilise means of economic outcome, the indices can reduce to a singular IOp measure that satisfies the *path independent decomposability*. As the indices satisfy the path independence axiom, then, according to Ramos and Van De Gaer (2012), the non-parametric direct and indirect (i.e. ex-ante (circumstances) and ex-post (efforts)) approaches provide

the same and reliable results[7]. This is crucial for the ex-ante approach used in this study, because it can thus be argued that efforts need not necessarily be observed.

Despite different inequality measures accepting the axiomatic properties above, they vary in sensitivity to different parts of the distribution (for a complete discussion on this refer to Chapter 2). Thus, also a **standardised** distribution version is computed, denoted by v_i^k.

In the non-parametric approach, hypothetical standardised outcomes are computed, denoted by v_i^k. To do so, first hypothetical mean outcomes with maximum inter-type inequality are computed following the commonly used approach by Elbers et al. (2008) and second, the hypothetical mean outcomes are standardised. The standardisation eliminates between type outcome inequality by replacing y_i^k with $y_i^k \frac{\mu}{\mu^k}$, where μ is the mean economic outcome of the entire sample. Here, the requirement for the path independence decomposability described above can be imposed, where $I(\{\mu_i^k\}) = I(\mathrm{y}) - I(\{v_i^k\})$. Foster and Shneyereov (2000) fully elaborate on this, but when the above axioms and the path independence property are satisfied and the inequality indices *I*( ) are restricted by using arithmetic means as the reference outcome, then the indices can be reduced to a single inequality measure (in this dissertation the MLD) denoted by E_0.

When accepting the path independence axiom on top of all the other axioms, FG (2011) argue that they can then further restrict the two scalar indices (2) and (3) to the following two unique non-parametric IOp indices (standardised and non-parametric):

A) Absolute Level of IOp (IOL): $$\theta_a = E_0\left(\left\{\mu_i^k\right\}\right) = \frac{1}{N}\sum_{i=1}^{N}\log\frac{\mu}{\mu_i^k} \qquad (4)$$

B) IOp Ratio (IOR): $$\theta_r = \frac{E_0\left(\left\{\mu_i^k\right\}\right)}{E_0(y)} \qquad (5)$$

In other words, the difference between the hypothetical and the actual outcome distributions are calculated, which gives the overall Inequality of Opportunity Level (IOL). When further divided by the MLD, it gives the non-parametric Inequality of Opportunity Ratio (IOR) value. The IOR can therefore be interpreted as the level (or share) of opportunity deprivation over overall inequality in outcome.

According to FG (2011, p 632), there are **three main advantages** in utilising these indices. First, they combine the ex-ante and ex-post approaches of IOp touched on before. Second, both indices satisfy the first four relative inequality index axiomatic properties and IOL satisfies two additional

[7] Roemer and Trannoy (2015) explain that the ex-ante approach is measures the inequality of the average outcome across types. And the ex-post approach is "obtained by rescaling the distribution of the outcome due to effort by the ratio of average income to average income per type". This is one of many options to control for the effect of circumstances on outcome. So, if it is possible to nullify the effect of efforts on outcome, then the ex-ante and the ex-post approaches will deliver the same results, under the path independence assumption. Thus, it can be argued that efforts need not necessarily be observed.

properties. Lastly, these indices are relatively easy to calculate and are identical to the Theil L subgroup decomposition. More information on the Theil L type decomposition, where types are exclusively categorised by circumstance variables, can be found in Haughton (2009, Ch.6, p.106 - 108).

As described before, the general IOp form is given by $y = f(C,E,u)$. By definition, circumstances are exogenous to the individual meaning that the individual cannot change these characteristics. However, one could argue that efforts, which are up to individual's choice, are affected by circumstances. This gives

$$y = f\left[C,E(C,v),u\right], \tag{6}$$

where *u* and *v* can account for numerous other factors such as luck or genetically inherited ability.

Since a) the ex-ante approach does not require to observe effort due to the path independence axiom, and b) this study accepts to only measure the effect or correlation of circumstances on economic outcome rather than inferring causality, the previous **form under (6) can be reduced** to:

$$y = \phi\left[C,\varepsilon\right], \tag{7}$$

where ϕ is a parameter and the whole equation is a log-linearised version of $\ln y = C\psi + \varepsilon$, where $\boldsymbol{\psi}$ includes all direct and indirect effects on outcomes (direct effect of *C* on *y*, indirect effect of *C* through *E* on *y*). This reduced-form (RF) can be estimated by ordinarily least square (OLS). From this, one can obtain a parametric estimate for the **smoothened distribution**:

$$\tilde{\mu}_i = \exp\left[C_i\hat{\psi}\right], \tag{8}$$

where $\tilde{\boldsymbol{\mu}}$ is a parametric analogue for the smoothened distribution μ_i^k and $\hat{\boldsymbol{\psi}}$ is the estimate parameter for the RF- OLS regression that contains all direct and indirect effects on outcomes. The predicted advantage is the same for all individuals with same circumstances. In other words, the distribution is simulated holding all circumstances constant.

$$\tilde{v}_i = \exp\left[\overline{C}_i\hat{\psi} + \hat{\varepsilon}_i\right], \tag{9}$$

gives the **standardised parametric** estimates, where $\tilde{v}$ is a parametric analogue of the standardised v_i^k and $\overline{C}$ is the mean circumstances across all observations. From there on, the parametric smoothened IOL and IOR indices can be obtained (smoothened and parametric):

Absolute Level of IOp (IOL): $\theta_a^p = E_0(\tilde{\mu})$ (10)

IOp Ratio (IOR): $\theta_r^p = \frac{E_0(\tilde{\mu})}{E_0(y)}$ (11)

And the standardised parametric estimates are as follows:

Absolute Level of IOp (IOL): $\theta_a^p = E_0(y) - E_0(\tilde{v})$ (12)

IOp Ratio (IOR): $\theta_r^p = 1 - \frac{E_0(\tilde{v})}{E_0(y)}$ (13)

In other words and similarly to the non-parametric approach, the difference in the inequality in the actual distribution and the simulated (standardised or smoothened) distribution gives the parametric IOL. The parametric IOR is this difference in relation to the overall inequality in outcome. In other words, the IOR (that is the relative level of inequality of opportunity) can be interpreted as the estimated percentage share of opportunity deprivation out of overall total inequality in outcome (that is the MLD).

FG (2011, p. 635) emphasise that 10 and 12, 11 and 13 do not give identical results in reality, despite the indices satisfying the path independence axiom. They argue that this is because they are estimated parametrically and using linear functional form assumption. Yet, the results they provide are very similar. The parametric IOp results FG (2011) provide in their paper are based on the estimation from (12) and (13). In this study also the results are very similar. The standardised distribution results are reported later.

Table 3.1 summarises the parametric and non-parametric IOL and IOR estimates for both the smoothened (within type elimination of different outcomes) and standardised (between type elimination of outcomes) distributions. As mentioned, FG (2011) use IOp estimation methods from (12) and (13) for the parametric approach and (4) and (5) for the non-parametric approach.

Table 3.1: Summary of Scalar Inequality of Opportunity Indices

	Parametric (standardised)	**Non-Parametric (standardised)**
IOL (main)	$\theta_a^{PS} = E_0(y) - E_0(\tilde{v}),$ (12)	$\theta_a = E_0\left(\left\{\mu_i^k\right\}\right) = \frac{1}{N}\sum_{i=1}^{N}\log\frac{\mu}{\mu_i^k},$ (4)
IOR (main)	$\theta_r^{PS} = 1 - \frac{E_0(\tilde{v})}{E_0(y)},$ (13)	$\theta_r = \frac{E_0\left(\left\{\mu_i^k\right\}\right)}{E_0(y)},$ (5)
	Parametric (smoothened)	**Non-Parametric (smoothened)**
IOL	$\theta_a^p = E_0(\tilde{\mu}),$ (10)	$\left(\theta_a = \mathrm{I}\left(\left\{\mu_i^k\right\}\right), \right)$ *(2)*
IOR	$\theta_r^p = \frac{E_0(\tilde{\mu})}{E_0(y)},$ (11)	$\left(\theta_r = \frac{\mathrm{I}\left(\left\{\mu_i^k\right\}\right)}{\mathrm{I}(y)}, \right)$ *(3)*

Source: Author's elaboration, summarised from FG (2011).
Note: IOL (main) and IOR (main) are the estimation equations reported in the empirical analyses sections.

In addition to the IOp indices, it is possible to compute the percentage share of the circumstances of IOR, that is the percentage contribution of each (statistically significant) circumstance of the opportunity deprived share of overall outcome inequality. FG (2011, p. 637) refer to this as the *partial IORs*. From the parametric RF-OLS regressions, it is possible to isolate the partial

effect of one (or a combination of) circumstances by **constructing a counterfactual distribution** following:

$$\tilde{v}_i^J = exp\left[\bar{C}_i^J \hat{\psi}^J + C_i^{j\neq J} \psi^{j\neq J} + \hat{\varepsilon}_i\right] \quad (14)$$

(14) is based on the standardised parametric form (9) and it is the counterfactual distribution holding one (or more) circumstances *J* constant, while all others are allowed to take their actual values. As a reminder, in (9) all circumstances were held constant (i.e. equalised) across all observations. The resulting counterfactual distribution from (14) allows for the computation of the circumstance *J* specific inequality shares (i.e. ***partial IORs***):

$$\theta_r^J = 1 - \frac{E_0(\tilde{v}^J)}{E_0(y)} \quad (15)$$

A few remarks on these partial IORs need to me made. First, only statistically significant (at the 10 percent level or higher) circumstances from the RF-OLS regressions are used to then compute the partial IORs. At each level of the empirical IOp analyses a new judgement is done on the inclusion of circumstances.

Second, there are several steps involved in computing the partial IORs meaning that the total of all computed partial IORs do not necessarily add up to 100 percent, explaining the total of the overall opportunity deprivation share of total outcome inequality.

Third, the partial IORs are only valid for circumstances. It is not possible to make a precise judgement on how much of the remaining outcome inequality is due to effort or other factors such as inherited skills, talents or luck.

Fourth, in order to compute the partial IORs and accept their results a series of assumptions need to be adhered to. These assumption are based on the validity and unbiasedness of the specific RF-OLS ψ parameter. The strong assumption is that omitted and unobserved variables from the RF are uncorrelated and orthogonal to the measured *C* in lny= *C*ψ + ε. Thus, FG (2011) argue that the partial IOR estimates are not lower bound estimates and should be interpreted only as the total impact of one (or more) circumstance(s) to IOR, i.e. the share of opportunity deprivation out of overall outcome inequality. Therefore, interpreting the partial IOR results must be done cautiously and only limited emphasis can be placed on them. One also needs to be cautious when using these results (even in combination with the RF-OLS regression results) for formulation policies. Nevertheless, they do give some general indication of the contributing amplitude of circumstances on the level of opportunity deprivation or enhancement and are complementary to the scalar IOp estimates and the RF-OLS regression results.

Fifth, in general the partial IOR estimates for consumption are expected to be higher than for income. This is because the IOR estimate serves as reference for the computation of the partial IORs. The IORs are lower for income because the residual inequality in the income distribution is considerably higher, which is consistent with the view that there is greater measurement error and transitory variance for income than consumption expenditure.

Last, the correct interpretation for partial IORs is important. The partial IORs are the maximum estimated shares that each circumstance has out of the overall level of inequality of outcome. This gives an indication to the extent of which the circumstance is an important element to boosting or hindering opportunity access. Is the RF-OLS coefficient estimate negative, the partial IOR can be interpreted as the share of aggravation of opportunity deprivation. Is the categorical circumstance seen

as enhancing opportunities (i.e. positive RF-OLS coefficient), the partial IOR is the highest estimated share of the circumstance to facilitate access to and boost opportunities.

A few **aspects for the indices' computation and result interpretations** later have to be taken note of. The first is the likelihood of omitted circumstance variables FG (2011, p.635 - 637). This calls to interpret the indices estimates, both IOR and IOL, as lower-bound estimates of IOp.

Second, the parametric estimates are expected to be lower than the non-parametric estimates and a mutual comparison between these results is essential in order to assure their validity. Furthermore, the inclusion of additional circumstance variables (unobserved, for instance) would increase the IOp estimates. This also suggests that the results estimated by (4), (5), (12) and (13) lower bound estimates.

Third, circumstance categorical variable division is kept to a maximum of three categorical variables in order to minimise inter-type variance, which is important for non-parametric analysis (FG, 2011, p. 640). This is essential because the non-parametric approach is sensitive to type conditional mean calculations.

Fourth, there is a chance of circumstance endogeneity. Gender, for example, as discussed later, can be viewed as partially endogenous to the individual. There may also be correlation, for instance, between circumstance variables, such as parental education being very low because they live/were born in rural areas where education accessibility is low or of inferior quality.

Last, there is always the possibility of selection bias (by data collector, researcher and person interviewed). For example, women may be more aware of HH consumption and characteristics than the men and are therefore the preferred individuals to be surveyed.

3.2.4. Complementarity Between the Parametric and Non-Parametric Approaches

There is a great amount of literature on the distribution free (i.e. non-parametric) and the distribution controlled for (i.e. parametric) analysis of (inequality) indices. This section only scratches the surface of this thorough analytical literature and aims at highlighting the most important aspects for the IOp analysis in this study. The comments and arguments put forward here are primarily taken from Slottje's (1990) paper, which discusses the trade-offs of both parametric and non-parametric approaches in constructing inequality in income indices. Cowell and Flachaire (2014) and Cowell (2000) also serve as reference.

Furthermore, Table 5.3 in the appendix summarises the main empirical IOp literature. It also contains information on which studies use a parametric, non-parametric or semi-parametric methodology. As can be seen, about half the empirical IOp studies use both parametric and non-parametric methodologies due to their complementarity in order to maximise information. This also goes for the direct ex-ante approach used in this study.

In general, many of the trade-offs depend on the quality of the data. Usually extensive micro or HH level datasets are used when trying to thoroughly analyse and decompose inequality. Certain dataset characteristics are crucial in determining which indices are the most suitable for analysing the data. Decisive characteristics include amongst others the size of the dataset and its distributional characteristics and representativeness. Chapter 2 gave a brief insight into the complexity of a) choosing suitable outcome variables, b) deciding on an optimal aggregation method for them and then c) choosing a suitable inequality index for the subsequent IOp analysis, bearing in mind that the inequality in outcome index has to satisfy a series of axiomatic properties that also match those of the IOp indices.

Slottje (1990) tries to reply to the question as to why functional forms should be imposed on data when computing inequality indices (i.e. using a parametric approach) when these indices can be

computed directly from the data without any great data conformity constraints (i.e. non-parametric approach). Unfortunately the answer is neither black nor white and especially when dealing with the relatively small dataset of this study with relatively low within type distribution (frequency), particularly then the two approaches are complementary. Both have their advantages and disadvantages. Slottje (1990) concludes that a) if only one method can be chosen then the parametric approach is preferred over the non-parametric. This is mainly because the statistical properties of the measures found by the parametric approach are superior to those of the non-parametric approach. The main reasons is that when an appropriate function form is used, the information obtained by a parametric analysis is richer (i.e. it maximises the information content): "you are filling in the shape of the distribution by imposing a functional form which gives you a multidimensional view of the level of inequality, which you can't get from a summary statistic directly".

b) When possible, both approaches should be used in order to check the results for consistency. Should the results be consistent, then it is likely that the data is of relatively good quality and that the model is suitable for it.

Table 3.2 summarises in brief some of the pros and cons for both the parametric and the non-parametric approaches when computing (inequality) indices, which also apply to the IOp indices of this study.

Table 3.2: Summary Trade-Offs Parametric Versus Non-Parametric Approach

	Parametric	Non-Parametric
Pros	- Imposition of functional forms, i.e. assumptions on the sample (e.g. normal distribution) - Can add information to the data by filling in the gaps. - In the IOp analysis of this study, hypothetical outcome distributions net of the effect of circumstances are computed. - Statistical inference relies on specific assumptions about the sample - Can provide good result when distributions are skewed and non-normal. Need for further investigation. - Can use when groups have differences in variability.	- Working directly with sample and flexibility, i.e. "distribution-free" methods (no functional form specifications) - Simplicity - Between and within type distributions analysis can be powerful (reliability of inference depends). - Assesses the median, advantage for some types. - OK to use for small group analysis, but large datasets preferred to boost result accuracy - OK to use for potentially non-normal distributed analysis - Less prone to be affected by outliers
Cons	- When the distributional assumptions are not reliable, it puts to question the results. - Some types in the IOp analyses have very small population sizes (in theory, minimum 15-25 observations per type preferred). - Analyses are affected by outliers, etc.	- Results can be greatly impacted by sample irregularities (e.g. bias), undermining any results inferences. - Groups should have the same or similar degree of dispersion.

Source: Author's elaboration summarised from Slottje's (1990), Cowell and Flachaire (2014) and Cowell (2000).

There are some general comments for the IOp indices' result expectations and interpretations of this study. First, the parametric estimates are expected to be lower than the non-parametric estimates, due to the imposition of functional forms (e.g. smoothing and standardising the distributions, computing hypothetical outcome distributions) for the parametric analysis.

Second, the circumstance categorical variable division is kept to a maximum of three categorical variables in order to minimise inter-type variance, which is important for non-parametric analysis (FG, 2011, p. 640). This is essential because the non-parametric approach is sensitive to type conditional mean calculations.

Third, a mutual comparison between these results is essential in order to assure their validity.

Last, The results interpretation sections (at all four levels of analyses) comments on possible divergences between the results of the two approaches.

This section (3.2) focused on the empirical IOp model used in this study and its adaptations to allow its operationalisation. To do so, first the research questions and hypotheses were stated. Second, an intuitive explanation of the model and the analysis methodology was given. Third, the econometric model for calculating the IOp scalar indices and singling out the circumstantial significance on economic outcome is stated particularly through the partial IOR approach. Last, this section

highlighted the complementarity between the parametric and non-parametric approaches used to compute the IOp indices. The section after this gives an insight into the country specific context, the NLSS III dataset, the dataset imperfections, the choice of independent and dependent variables and the econometric model adaptations to the them.

3.3. Inequality of Opportunity in Nepal: Country Specific Background, Data, Variables and Model Adaptations

This section first provides some general background and context information on Nepal, its socio-economic, political and cultural context and geographical conditions. Second, details on the characteristics of the dataset, its shortcomings and how they are dealt with, and the dataset constraints that are applied for the IOp analyses are provided. Third, this section elaborates on the dependent variables, their definitions and their particularities (main reference remains Chapter 2). Fourth, an investigation is done into the choice of circumstance variables and how they were adjusted for this study. Last, an overview of the main econometric model adaptations to the NLSS III dataset and the Nepalese country specific context is given.

3.3.1. Country Specific Background and Geographical Scope

This study is primarily an empirical analysis of inequality in outcome and IOp with an exclusive focus on Nepal. In order to be able to analyse the Nepalese dataset and interpret the results of this study with the upmost care, it is crucial to have a rough understanding of the cultural, historical, socio-economic and geographical context of Nepal.

Unless otherwise stated, the facts and figures referred to in this section are mainly taken from the World Bank's World Development Indicators (WDI, 2012), the Nepalese Central Bureau of Statistics (2012) and the CIA World Factbook (2012 and 2013).

Nepal ranks as one of the 30 poorest countries in the world according to the WDI (2012). The World Bank (2016, p. 4) states that "Nepal is at a very early stage of development and that the engine of growth has not started cranking up to the extent where this growth acceleration would begin to generate a widening consumption distribution". Its **population** totals about 26.5 million, 25% of which live below the poverty line and its demographic growth rates are over 1.35% per year (CBS, 2012). Its **GDP growth rate** in 2012 was 4.63% and having 3 main income sources. First, about 25% of Nepal,s GDP come from foreign aid, about 1,581 million USD (ADB, 2018; Ministry of Finance, 2017). Second, remittances also make up about 25% of Nepal's GDP (ADB, 2018; Ministry of Finance, 2017). Last, the agricultural sector contributed about 37.1% to Nepal's GDP in 2011, however this share has been falling rapidly over the past few years to as low as 27.6% in 2017 (Ministry of Finance, 2018). Inflation rates have been hovering between 4% to 5% for the past few years. The country's Gini Coefficient is estimated to be relatively low at 0.328 in 2011 (compared to 0.414 in 2003/2004) and about 65.9% of the population was literate (75.1% of men and 57.4% of women). Bhattachan (2003, p. 5) claims that "Nepal is a rich country inhabited by poor people". It is a socially diverse, rich in water resources and biodiverse country, but it is "craving for irrigation and electricity".

In 2000, the Nepalese government initiated the creation of Special Economic Zones (SEZs) across the country in order to boost the dynamics, diversification and competitiveness of its economic structure. The SEZs are free trade zones that are ought to encourage foreign direct investment (FDI), develop export capacity productions, and encourage overall infrastructure development within the regions where the SEZs lie (Special Economic Zone Committee, 2017). Figure 3.8 illustrates the locations of some of the SEZs across the country and their respective development stages.

Nepal's **political reality** is rather peculiar. At the domestic front, the past five decades were marked by great turmoil. The country has seen anything from riots, revolutions, ethnic cleansing, refugee crises, monarchical overthrow and democratic restructuration. Despite the establishment of a multiparty democracy in 1990, which was based on a constitutional monarchy, Nepal has been confronted by severe Maoist uprisings and extremist movements ever since. In 2008, Nepal was declared to be a federal democratic republic, its monarchy was abolished and its government coalitions changed four times in the following three years. Its newest constitution was passed in 2015, which ever since has led the country into an uncertain and yet structurally determined federalisation process. This process as by fall 2019 is still in progress and continues to create much political and public tensions.

At the international political front, Nepal is caught in geopolitical tensions between bordering India and China. Much of these tensions arise due to Nepal's wealth in natural resources such as being at the source of clean Himalayan water and the Tibetan refugee crisis.

Nepal is a **culturally rich** country with over 126 different indigenous, ethnic and caste groups, over 123 different spoken languages and more than 10 well represented religious groups (CBS, 2012). The country's socio-cultural structure is marked by its Hindu originated caste system, which came from India in the 11th century (Bhattachan, 2003).

Being landlocked between China and India, Nepal is a small nation in South Asia. Its land surface is comparable to that of Greece or Tunisia (CIA World Factbook, 2013) and it is marked by extreme topography. Nepal is famous for being a hikers paradise, because it is home to the top of the world (Mount Everest) and seven of the world's "8-thousanders".

On an **ecological level**, the country is divided into three west to east belts (see Figure 3.9). First, the northern mountain belt, which bears the high Himalayan mountain range and is marked by harsh climates. Second, the central hill belt, which is marked by lower altitude rolling hills and picturesque lakes. This belt also hosts the country's capital Kathmandu and other major cities such as Pokhara and Lalitpur. Third, the Terai belt, which is famed for its fertile agricultural low lands, the countries lush forests and nature reserves.

For administrative purposes, the country is divided into five **Development Regions**, which are sub-divided into a total of 75 districts, also indicated by Figure 3.9. These are: the far-western, the mid-western, the western, the central and the eastern regions.

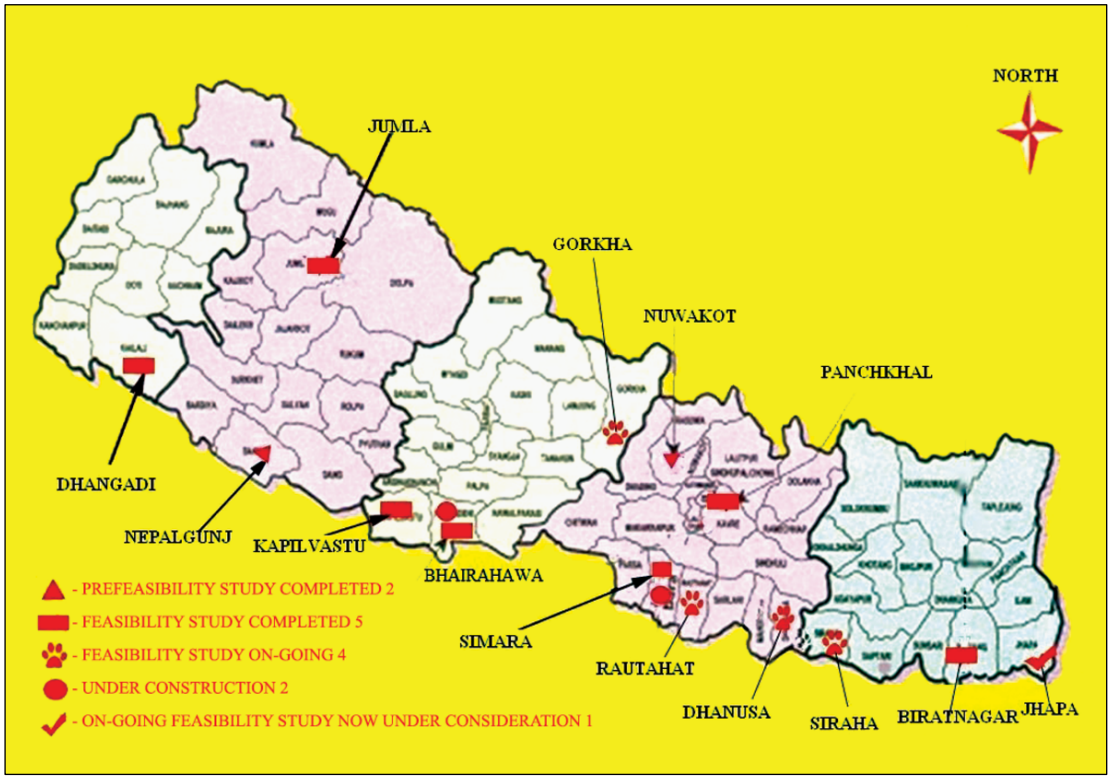
Today Nepal is a largely centralised country in economic, political and social terms. Kathmandu serves as the hub for almost everything. Yet historically, the country's capitals (royal, imperial and feudal) have changed locations over the centuries, many of them still being urban centres today. Figure 3.10 shows the dispersion of **urban centres** across the country with rough population estimates (Fafchamps and Shilpi, 2003).

Kathmandu serves as the political and administrative hub hosting virtually all national public institutions. It is the country's economic centre in terms of economic production and output, and it is the labour market centre hosting most the country's few job opportunities in public institutions, the service sector (other than tourism), non-governmental and international organisation, and politics.

Furthermore, it is important to add that Nepal is highly susceptible to a vast range of **natural disasters** due to the divergence of extreme topography and climatic changes on its relatively small territory. These include, for instance, flooding, earthquakes and landslides. Nepal was ranked as the 11th most earthquake vulnerable country in the world (ADB, 2018). Furthermore, Nepal's low level of development and largely impoverished population also makes them vulnerable to additional disasters such as epidemics.

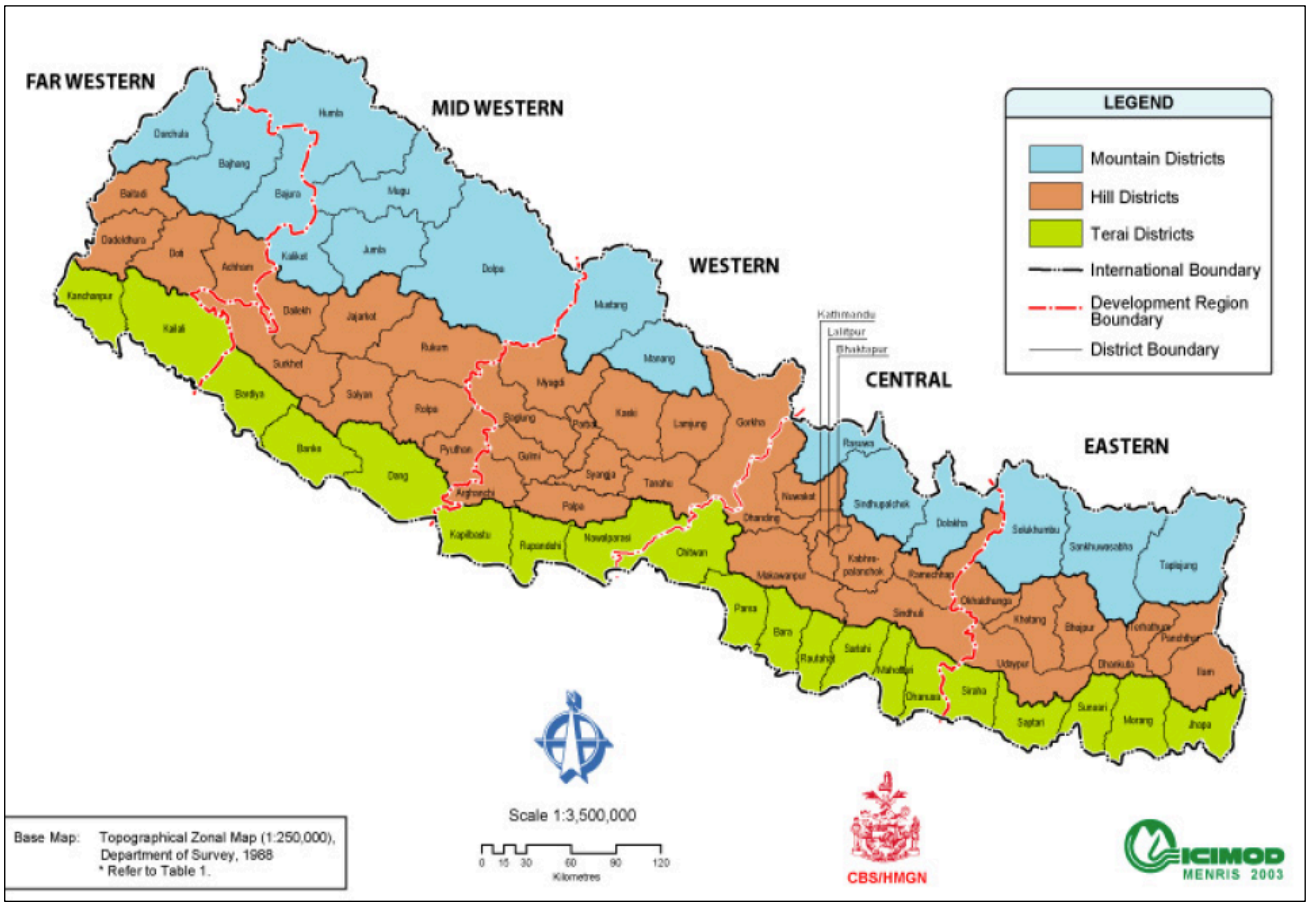
On the whole, Nepal is a slowly emerging country, rich in resources, biodiversity and cultures, that has suffered from decades of political instability, which has failed to created social cohesion and economic prosperity, so that over a quarter of its population still lives below the poverty line today.

Figure 3.8: Map of Nepal's Special Economic Zones



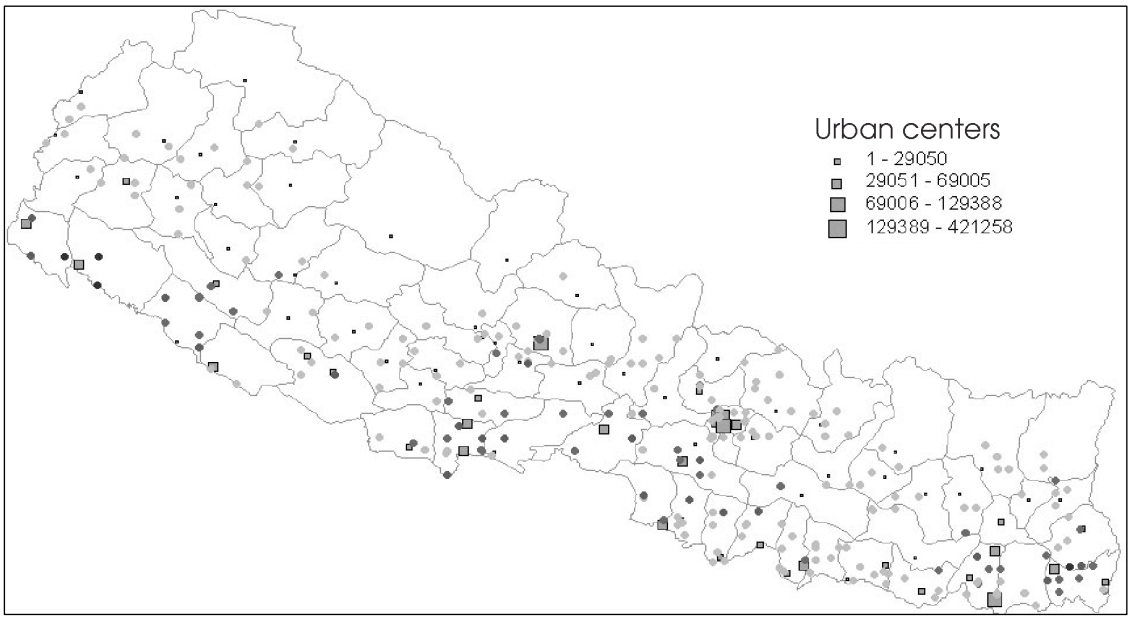
Source: Special Economic Zone Committee (2017)

Figure 3.9: Administrative Boundaries of Nepal



Source: Nepal Central Bureau of Statistics, International Centre for Integrated Mountain Development (2003).

Figure 3.10: Urban and Rural Areas



Source: Fafchamps and Shilpi (2003)

3.3.2. Data, Data Cleaning, Data Constraints, Variable Definitions and Adaptations

3.3.2.1. Data

This section provides some additional information on the NLSS III dataset to what was already described in Section 2.3 on the data description for the dependant variable aggregation methodologies and the study of inequality in outcome.

This dissertation uses the Nepal Living Standards Survey (NLSS) III. The NLSS III is the third nation-wide household survey conducted from February 2010 through to February 2011. Table 3.3 provides some descriptive details on all three NLSS dataset rounds. The number of observed HHs and individuals increased each round, yet the panel sample is very small. It appears that there are disagreements on the panel data, its roaster information and the panel sampling and weighing methodology, thus the panel identification specifications are not officially provided when purchasing the NLSS data. For further details, contact the CBS of Nepal.

The data collection for a fourth round, NLSS IV is being done at the moment of writing (2019/20) and the dataset is supposed to be made available for purchase from fall 2020 (World Bank, April 2019). Limited official details have been released on the data, but the author was told informally by CBS officials that it contains more and richer information particularly on the socio-economic and cultural heritage elements of individuals. Also, the dataset will contain over 9,000 HH in over 750 primary sampling units across the country.

Table 3.3: Details of the Nepal Living Standards Surveys (NLSS)

	NLSS III	**NLSS II**	**NLSS I**
Survey	Third Nepal Living Standards Survey (NLSS III)	Second Nepal Living Standards Survey (NLSS II)	First Nepal Living Standards Survey (NLSS I)
Surveyed Year	Feb. 2010 – Feb. 2011	April 2003 – April 2004	1995 -1996
Survey and Reports published	2011	2004	1996
Households surveyed Cross-section sample (Panel)	5,988 (1,032)	3,912 (1,160)	3,373 (-)
Individuals	34,146	21,531	18,855

Source: CBS (2011)

3.3.2.2. Dataset Sample Constraints for the Inequality of Opportunity Analyses

Table 3.4 summarises the NLSS III dataset characteristics with and without data constraints. Table 3.5 gives some descriptive statistics for the HH size at the national level and with various dataset constraints. The original dataset counts over 5,988 HHs and over 34,146 individuals. FG (2011) propose several data constraints in order to allow for great results comparability in their own study of six different Latin American countries as well as to other empirical IOp studies. In order to enhance the comparability of this study's results the same **3 data restrictions** are applied to the NLSS III dataset. Additionally, one of them, namely HH composition, is slacked and additional results are provided. This expands the analysed sample size by over 1,150 observation and renders the analysis more robust.

The data sample constraints are:

1. Age: 30 – 49 years
2. HH members:
    A. All HH members
    B. HH head and spouse
    C. HH head/ spouse, sons/daughters, brothers /sisters
3. Complete information on all circumstance variables

First, observations must be **aged between 30 to 49 years**. This was done, because individuals in this age cohort are considered to be the most economically active and independent. When applying this restriction, only 7,119 individuals remain out of the original 34,146. Note that this restriction criterion could be slackened to include, for example, individuals aged above 18 years in order to have a larger sample. This however is not done in this study in order to allow for comparability to other studies.

Second, FG (2011) propose that the analysed individuals must be either the **HH head or the spouse.** Figure 3.11 illustrates the proportion of individuals by relationship to HH head and age. By far the largest proportion is that of HH heads and spouses, followed by sons and daughters as well as sons and daughters-in-law. The first reason as to why FG (2011) applied the HH head and spouse restriction was to be able to cross-compare the results of the six Latin American countries they analysed. Two of these countries' datasets, Brazil and Peru, only contained information on HH heads and spouses, while the others contained information for all HH members. The second reason for only using HH heads and spouses in the data sample is, because they are deemed to be the age cohort with the highest proportion of employed individuals, therefore having representative information on the outcome variables. Here, about 38% of observations (i.e. 11,615 individuals) in the total dataset are either a HH head or spouse. The HH head and spouse data constraint is only applied and reported for the national level analyses in Section 3.4. Additionally, two other dataset HH composition variations are analysed. The first includes all HH members (i.e. the complete 34,146 individual dataset). This bears a few advantages. It allows to compare this study's results to other IOp studies except from FG(2011) and it increases the sample size analysed. The second includes HH head and spouses, sons and daughters and brothers and sisters to the HH head (22,331 observations in the NLSS III dataset). The extension to brothers and sisters, and sons and daughters was done, because during the data cleaning process, parental education information could only be filled in for these HH members. So, it serves as a pure robustness check to the HH head and spouse sample and the all HH member dataset sample. Refer to Section 3.3.2.5.2 for more details on the filling in of parental education background.

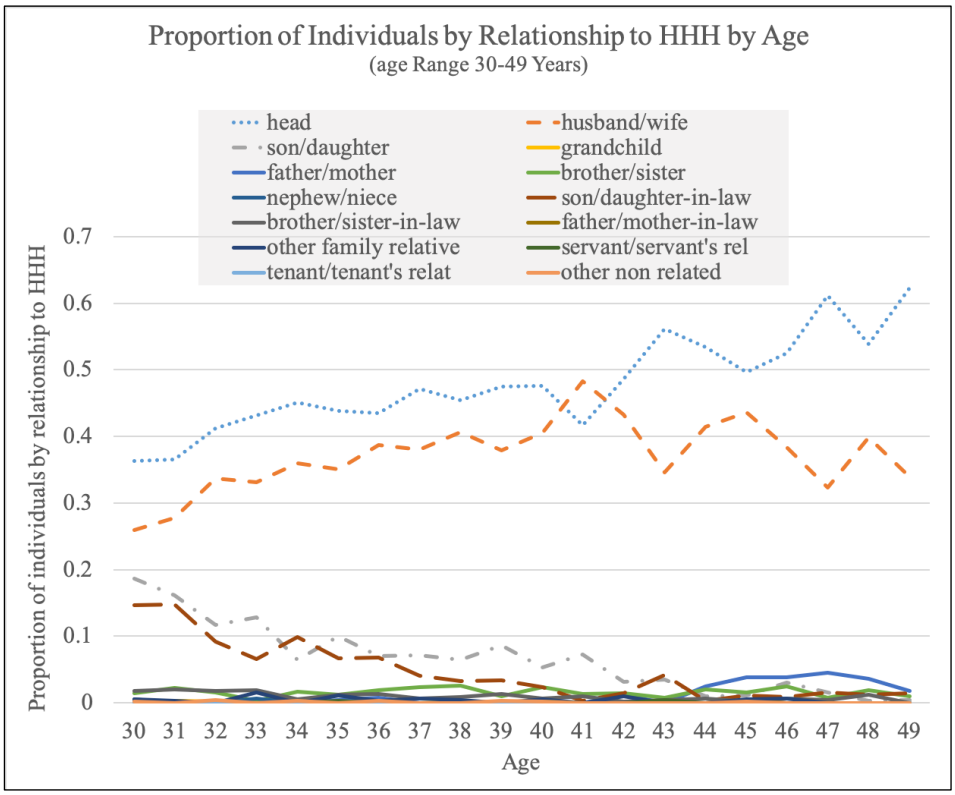
Last, all observations must have **complete information on all circumstance variables**. The dropped observations, on top of the first two data restriction criteria, where almost exclusively individuals who had lacking information on the birth region 1 (urban-rural area) circumstance variable.

Table 3.4 shows the final dataset characteristics when all dataset constraints are applied. These are applied to both consumption and income. Overall, for all HH member constraint variations and for both dependent variables, this study uses only between 17% to 21% of the overall dataset for the IOp analyses. For instance, the data sample for consumption expenditure for all HH members aged between 30 to 49 years and with complete information on all circumstances had a total of 7,044 observations, i.e. 21% of the total dataset. Since various dataset constraint were applied and different assumptions applied, the sample size decreased, meaning that extreme values and larger type variance will impact on the econometric analyses. Hence, one must remember this when interpreting this study's results.

Since the NLSS III data was reweighted in order to be nationally representative and this study only uses a sub-sample of it, precaution must be taken when interpreting the results. Furthermore, the NLSS III raw dataset was adjusted for regional cost of living, meaning that this study did not need to readjust for regional living cost. An overview of the importance and procedure for adjusting regional cost of living divergences can be found in Deaton and Zaidi (2002, p. 40-45).

Table 3.5 shows that the applied data constrains have various impacts on the HH size. This is important for some of the latter descriptive result interpretations and the judgments on HH welfare.

Figure 3.11: Proportion of Individuals by Relationship to HHH by Age



Source: Author's elaboration. *Note*: Age restriction 30-49 years.

Table 3.4: NLSS III Data Summary With Data Sample Constraints

Households surveyed	5,988	
Individuals	34,146	
Number of obs. (share of original sample, %) 1) aged between 30 to 49 years	7,119 (20.8)	
Number of obs. (share of original sample, %) 2) A) all HH members B) household head or spouse, or C) HHH/spouse, sons/daughters, bros/sisters	C) 34,146 (100) D) 11,615 (30.0) E) 22,331 (65.4)	
Number of obs. (share of original sample, %) 1) aged between 30 to 49 years, and	Per capita consumption expenditure	Per capita income
2) A) all HH members B) household head or spouse, or C) HHH/spouse, sons/daughters, bros/sisters 3) who have information on all circumstance variables out of all survey obs.	A) 7,044 (20.6) B) 5,899 (17.3) C) 6,530 (19.1)	D) 7,106 (20.8) E) 5,958 (17.4) F) 6,623 (19.4)

Source: Author's computations.

Table 3.5: HH Size Descriptive Stats According to Dataset Constraints, National Level

NLSS III Dataset Sample	Observations	Mean Household size	Std. Dev.	Min.	Max.
Complete dataset	5,988	4.79	2.33	1	21
With data constraints 1) all HH members, 2) aged 30 - 49 years, 3) information on all circumstances	6,974	5.41	2.54	1	21
With data constraints 1) HH head and spouse, 2) aged 30 - 49 years, 3) information on all circumstances	5,849	4.99	2.01	1	21
With data constraints 1) HH head and spouse, sons and daughters, brothers and sisters, 2) aged 30 - 49 years, 3) information on all circumstances	6,505	5.24	2.34	1	21

Source: Author's computations. *Note*: The slight divergence in number of observations compared to Table 3.4 is due to the fact that Table 3.4 reports the observations with the mentioned 3 dataset constraints, whereas here, individuals must also have information on both of the dependent variables.

3.3.2.3. Dataset Shortcomings Explained

There are some dataset imperfections that can be generalised across most LSMS datasets. Much literature discusses the design, collection, representativeness and analysis of HH surveys (e.g. Deaton, 1997; Grosh and Glewwe, 1998 and 2000). Some general traits and shortcomings for HH survey data include, first, the trade-off between gathering a maximum amount of information from individuals and

the quality of information received. Second, the trade-off between the number of observations and representativeness of the data with the exhaustiveness of it. Third, the periodic distance between data collection rounds due to the need of great resources to conduct the surveys. For instance, Nepal collects annual population and housing census data, while the NLSS rounds are more than 7 years apart.

Furuta (2016) highlights three major general data shortcomings in the NLSS dataset rounds. First, he regrets the exclusion of institutional HHs. Second, local measuring units were used, which can increase recording error and can make the utilisation of the data difficult, despite the existence of a conversion table. The author also noted some inconsistent and implausible values given, when cleaning the data, particularly during the aggregation process of the dependent variables. Third, the data was recorded based on the memory of the respondents. For example, individuals were asked how much rice they purchased over the past week, or the past month, or the past year. As elaborated in the dependent variable aggregation sections, the author also found significant discrepancies between the given replies.

Furthermore, the CBS (2011) NLSS III reports highlight that a specifically developed software for the NLSS III survey was used for the data collection. However, some challenges using the software were encountered possibly inducing data entry errors. The CBS guarantees though that maximum effort was done to adequately train interviewers in order to minimise error and bias.

During the data cleaning and descriptive analysis of the NLSS III dataset, the author found three additional data shortcomings that are worth mentioning in brief (additional details can be obtained upon request). Also, the attempts to overcome these imperfections are discussed. These three shortcomings are and follow in the next three sub-sections:

A) High proportion of absentees/migrants
B) HH size divergence
C) Disproportionately low sex ratio

3.3.2.3.1 Data Shortcomings: High Proportion of Absentees/Migrants

A high proportion of the Nepalese population is absent from the HH they originally belong to. These individuals are classified as absentees[8] and information on them is collected in a separate section of the NLSS survey with less complete information on socio-economic elements. However, they cannot be included directly in the IOp analysis, since they do not have complete information on key IOp circumstance variables such as parental education, religion, or ethnicity and caste group belonging. Table 3.6 compares the existence of circumstantial variables and possible inferences from the observations present in Nepal to those captured as absentees in the absentee section of the NLSS survey. Overall, it was judged that absentees (abroad) cannot be included in this IOp study since they lack crucial information on circumstance variables. Nevertheless, one can argue that this is not biasing the rigor of this study, because it tries to analyse the level of opportunity deprivation (and inequality of economic outcomes) of the people within Nepal.

The following statistics come from the CBS (2011 and 2012), the Ministry of Labour and Employment (2014 and 2018) and the World bank (2011). 37% of the total Nepalese population has migrated from other places (VDC, municipality or outside the country) to current place of residence

[8] CBS (2011, Report 2, p. 132): "Absentees and persons who have left one household to live permanently elsewhere would be enumerated in their new location if they satisfied the 6 months criterion in their new location. Consequently, there is no conceptual omission or duplication in determining the population within Nepal. Absentees who have moved outside Nepal are excluded from the Nepal population."

(this includes both, migrated individuals captured in the survey in a new place of residence and those completely absent only captured in the absentee section of the survey). 20 to 25% of the total population (domestic and international, all ages) is found to be absent or away from home. This captures only individuals marked as absentees. 55 % of the total absentee population is currently residing within Nepal while the remaining 45 % resides outside Nepal (of which over 20% in India and are thus not captured at all due to the open boarder). More than 84.9% of absentees are in the working age population (15-59 years, NLSS III definition of working population). The gender gap between male and female absentees is striking but not surprising. More than 70% of all absentees are male. Figure 3.12 illustrates the relative propensity to migrate by relationship to HH head. It illustrates clearly that the category of sons and daughters as well as husbands and wives are the largest in number, and of course more so for the male part of these categories.

Furthermore, 53% of HHs have at least one absentee currently living either abroad or within the country. Some HHs have both, at least one person living abroad and at least one person being absent but living within the country. Their **remittances** (in liquidity or in-kind) are seen as an important contributor to Nepal's economy. The MLE Report (2014), for instance, suggests that migrant remittances contributed an ever increase share of GDP and as much as 29.1% in FY2013/2014 (76% of which are international remittances). This is arguably an important element impacting on the inequality and opportunities for those who remain at home. Section 3.10.2 briefly goes into an IOp analysis extension idea by which to include the influence of parental presence and remittances.

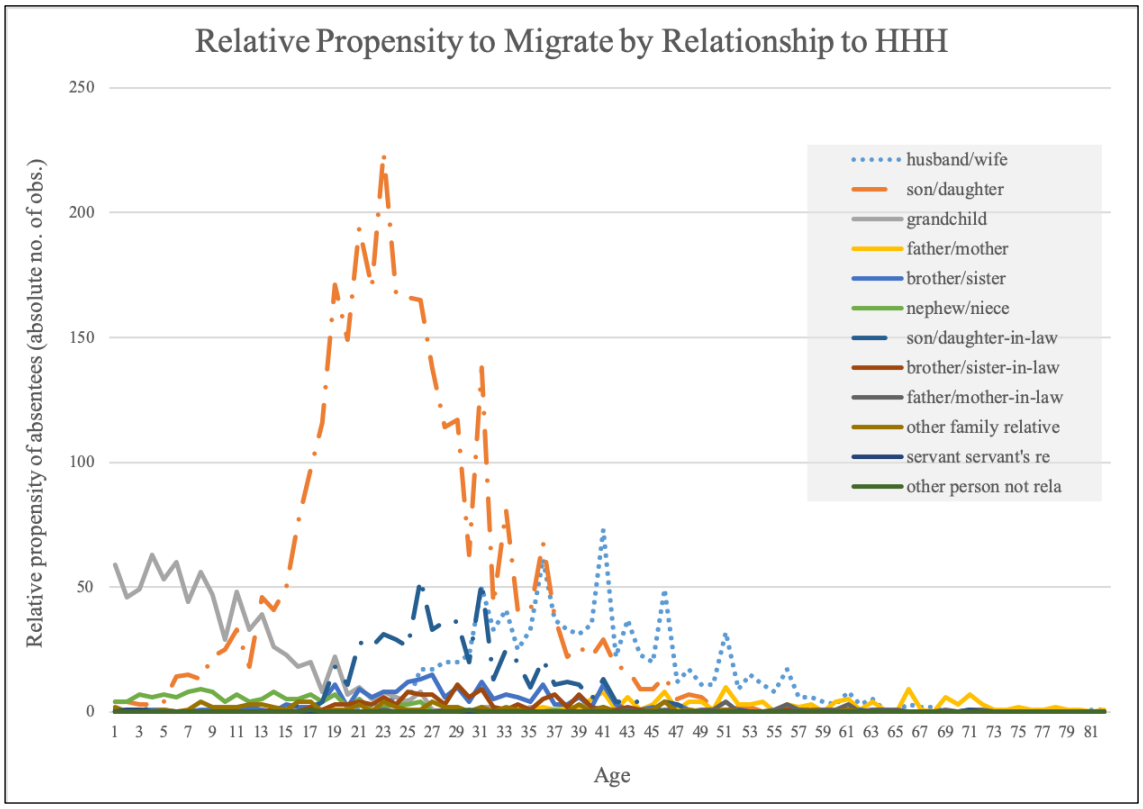
Emigration has two main important implications. First, concerning the IOP analysis where emigration is seen as both an exogenous and endogenous (effort) variable. According to the author's analyses and interpretation of the absentee descriptive statistics, it can be concluded that a) most emigrants migrate after the age of 18 and b) people with at least some basic educating have a higher propensity to emigrate. This suggests that educated emigrants are bigger risk takers than non-emigrants. Furthermore, emigration has (aspired to) economic welfare benefits for the remaining HH, thus receiving higher education in Nepal can be seen as an opportunity-equalising tool since it increases the individual's propensity to migrate.

Second, concerning the data analysis, where the remaining population is different to the emigrating population in terms of educational background. Since emigration is in part an endogenous choice, excluding absentees from the analysis should overestimate IOp.

Furthermore, an effort was made to investigate into the possible inclusion of absentees into the IOp analyses. Table 3.6 compares the existence of circumstantial variables and possible inferences from the observations present in Nepal to those captured as absentees in the absentee section of the NLSS survey. As can be seen from the table, however, is that absentee observations severely lack information on circumstances for them to be included in the IOp analyses. Therefore, when interpreting this study's results, care must be taken acknowledging that a substantial part of the more educated, risk neutral and entrepreneurial of the population is not included in the analyses.

Overall, it is important to acknowledge that the sample population used in this study's IOp analyses is not fully representative of the entire population, because a) the sex ratio is significantly lower than the national average (see Section 3.3.2.3.3), and b) absentees are excluded (they benefit from migrational opportunities and are arguably different to the remaining population).

Figure 3.12: Relative Propensity to Migrate by Relationship to HHH



Source: Author's elaboration.

Table 3.6: Absentee Circumstantial Variable Comparison

Variables used for the IOp analysis in the nationally representative dataset	**Variable in the dataset for absentees**	**Indirectly included variables for absentees**	**Not included variables for absentees**
Circumstances included • Gender • Parental education (mother's & father's education) • Religion • Ethnic and caste group belonging • Birth region 1 (Urban / Rural) • Birth Region (Mountain / Hill / Terai) Circumstances deemed as crucial but not included • Parental occupation	Circumstances used for the nationally representative dataset • Gender Other variables that could, to some extent, be of interest for the IOp analysis • Age, highest educational attainment, living location (country and rural/urban), migration reason, type of job, main primary activity now, remittance information (how often remittances received, amount, means of remittance, in kind of liquidity, usage of remittance, etc),…	• Parental education → indirectly through relationship to HHH educational attainment → filled in from generation to generation. Difficult for some generations. • Religion → indirectly through relationship to HHH, but religion only noted for HHH → hereditary assumption's feasibility is questionable. • Ethnicity and caste group belonging → indirectly though relationship to HHH because hereditary. Strong inter-caste mobility assumption • Birth region 1 & 2: maybe indirectly through relationship to HHH, but strong immobility assumption	• Parental occupation (does also not figure for remaining population) → too strong assumptions needed to fill this in (see Section 3.3.2.5.3).

Source: Author's comparison.

3.3.2.3.2 Data Shortcoming: HH Size Divergence

Figure 5.3 in the appendix shows the percentage distribution of HHs by size and their evolution across all three NLSS dataset rounds. It becomes evident that the share of HHs with less than 5 people has increased and the share of HHs with more than 4 people has decreased since the first NLSS in 1995/96. Table 3.7 shows the distribution of HH size using the NLSS III dataset. Over 50% of HHs have 4 or less HH members. Less than 6% of all HHs have more than 8 members.

According the Population and Housing Census (CBS, 2012) the average nation-wide HH size has decreased from 5.44 in 2001 to 4.88 in 2011. The NLSS III reports (CBS, 2011) talk of an average HH size of 4.9, which stands marginally in contrast with the population census results for the same year, and what the author computed using NLSS III dataset (4.8; see Table 3.8). Table 3.8 displays some descriptive statistics for HH size depending on the applied IOp population sample constraints. Some variance can be noticed. For instance, when only HH heads and spouses are considered, aged 30 to 49 years with complete information on all circumstances, then the HH size is 4.99. However, when all HH members are considered, it rises to 5.41. This, however, does not directly impact on the IOp analyses, because the per capita economic outcome variables are computed before the dataset constraints are applies. This means that a similar HH size distribution and average to the CBS (2011) results is applied in the IOp analyses.

Table 3.7: HH Sizes

Number of people living in the HH	**Freq.**	**Percent**	**Cum.**
1	204	3.41	3.41
2	638	10.65	14.06
3	923	15.41	29.48
4	1,259	21.03	50.50
5	1,083	18.09	68.59
6	795	13.28	81.86
7	469	7.83	89.70
8	267	4.46	94.15
9	133	2.22	96.38
10	84	1.40	97.78
11	49	0.82	98.60
12	25	0.42	99.01
13	17	0.28	99.30
14	9	0.15	99.45
15	11	0.18	99.63
16	6	0.10	99.73
17	5	0.08	99.82
18	2	0.03	99.85
19	5	0.08	99.93
20	3	0.05	99.98
21	1	0.02	100.00
Total	**5,988**	**100**	

Source: Author's computations. *Note*: using the NLSS III dataset at the national level.

Table 3.8: HH Size Descriptive Stats According to Dataset Constraints, National Level

NLSS III Dataset Sample	Observations	Mean Household size	Std. Dev.	Min.	Max.
Complete dataset	5,988	4.79	2.33	1	21
With data constraints 1) all HH members, 2) aged 30 - 49 years, 3) information on all circumstances	6,974	5.41	2.54	1	21
With data constraints 1) HH head and spouse, 2) aged 30 - 49 years, 3) information on all circumstances	5,849	4.99	2.01	1	21
With data constraints 1) HH head and spouse, sons and daughters, brothers and sisters, 2) aged 30 - 49 years, 3) information on all circumstances	6,505	5.24	2.34	1	21

Source: Author's computations. *Note*: The slight divergence in number of observations compared to Table 3.4 is due to the fact that Table 3.4 reports the observations with the mentioned 3 dataset constraints, whereas here, individuals must also have information on both of the dependent variables.

3.3.2.3.3 Data Shortcoming: Disproportionately Low Sex Ratio

When investigating into the NLSS III data, the author noted an extremely low sex ratio, i.e. the overrepresentation of women and a much too small male population. The sex ratio is defined as the proportion of man compared to women, i.e. number of men in relation to 100 women (WDI, 2012). Table 3.9 gives some general sex ratio descriptive statistics and Table 3.15 displays the sex ratios for the national level IOp analyses depending on the different population samples analysed. Figure 5.2 in the Appendix shows the evolution of the sex ratio for all three NLSS rounds. Appendix 4: provides additional details to the sex ratio discussion that follows.

The sex ratio in this study is disproportionately low (78 for the national analyses with all three data constraints for all HH members) compared to the overall national representative NLSS III dataset level (85) and the national population census for 2011 (92). The NLSS III Report (CBS, 2011) only describes, but does not interpret or explain the low sex ratio in the dataset. Note that most of the IOp literature does not report the sex ratio in their descriptive statistics. The most extreme case found was in the FG (2011, p.642) paper, where sex ratio drops to low as 81 for Colombia. This is still not as low at the sex ratio for the sub-sample in this study, however, the FG (2011) paper does not address the sex ratio issue.

In order for this study to be academically viable, the utilised dataset must be representative of the national population. The key is to investigate into the degree of representativeness of the utilised sub-samples in this analysis vis-à-vis the total survey. The following tries to investigate into a) why the sex ratio drops so drastically in the sub-samples, and b) the randomness of the response rate of individuals in the sub-samples.

The author has isolated three main and three minor factors that are seen to disproportionately lower the sex ratio of this study. First, the three minor factors are briefly discussed.

1. Data constraints for IOp analyses

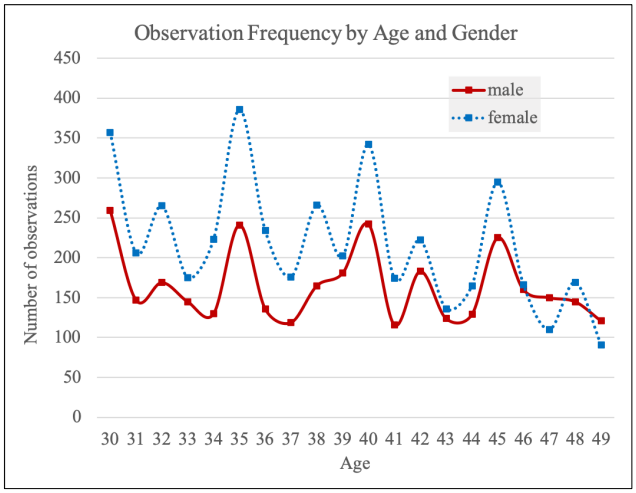
This study applies three dataset constraints for the IOp analyses (discussed in Section 3.3.2.2) in order to make its results comparable to those by FG (2011). Figure 3.13 and Figure 3.14 illustrate the absolute and relative gender observation frequency for the 30 to 49 years age range. They reveal a

relative skewness of male observations towards the older age cohorts suggesting that the frequency of men reporting is skewed towards the older age groups in relative terms. This may be due to the fact that a large number of men are absent, i.e. they are international labour emigrants.

Furthermore, there is an obvious gender gap in relative terms, where women represent a relatively larger proportion of respondents and that particularly in the 30 to 49 year age cohort. However, this gender gap decreases as one moves along the x-axis towards the older ages. In relative terms, the number of men reporting increases towards the older age cohorts. The gender gap is reversed after the age of 45 years. Equally so, women report more, in relative terms, when looking at the 30 to 45 years age range. This also could underpin the large absence of male offspring (absentees) due to domestic and particularly international migration.

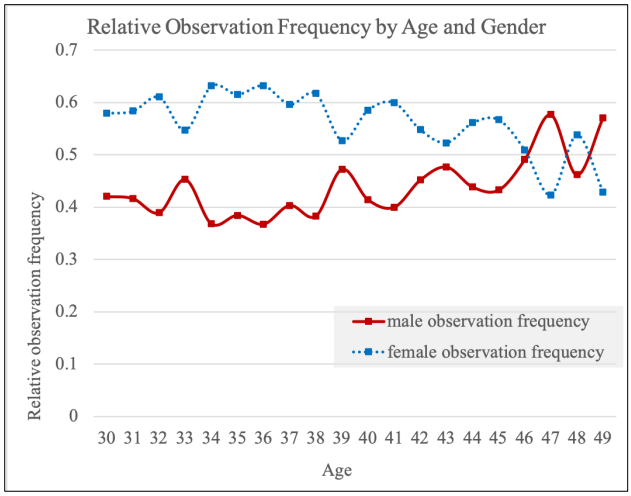
There are 5-year spikes, especially visible in diagram 1. Such spikes can be seen as normal, due to age being self-reported in this survey (NLSS III HH Survey; CBS, 2011, p.81, "How old is …NAME…? Age in completed years"). Ironically, the spikes in diagram 1 are more pronounced for women than for men but are slightly evened out for both sexes when looking at the relative age and gender distribution diagram. However, diagram 2 still shows some unknown spikes, for example at ages 33, 39, 43 and 47.

Figure 3.13: Absolute Observation Frequency by Gender and Age



Source: Author's computation
Note: Age is restrained to 30 to 49 year olds. Relative observation frequency is the absolute number of observations by gender over the total amount of observations for that age group.

Figure 3.14: Relative Observation Frequency by Gender and Age



Source: Author's computation
Note: Age is restrained to 30 to 49 year olds. Relative observation frequency is the absolute number of observations by gender over the total amount of observations for that age group.

2. Sex Ratio at Last Birth (SRLB)

Traditionally, Nepal is a country where male offspring are preferred over female offspring for socio-economic and cultural reasons. This has caused waves of infanticide and illegal abortions in Nepal. Abortions were legalised in Nepal in 2002 and Frost et al. (2013) elaborate on what they refer to as Sex Ratio at Last Birth (SRLB). This means that parents, unknowingly of the gender before birth and before the legalization of abortions, would have their first child. If it is a boy, they stop having more children unless their economic situation would allow it. If it is a girl, however, the parents would "try" for a second child. If it is a girl, again they would "try" to have a boy. This process continues until either the family has a boy, or the economic situation of the family becomes too unbearable. This

means, however, that there are countless families where there is a correlation between the number of children and the increase in girls, causing a slight increase in the sex ratio.

Yet, Frost et al. (2013) argue that before (and after) the legalization of abortion in Nepal there is no evidence of sex-selective abortion. Changes in the sex ratio at birth since the legalization would suggest an association with it, despite sex-selection being explicitly prohibited.

So, the extent to which this factor increases the overrepresentation of women in Nepal, is uncertain.

3. War, civil conflict

For the past century, Nepal was haunted by countless wars and civil conflicts. These have a tendency to decimate the number of men, who are the most prone to partake in the conflicts. Quy-Toan and Iyer (2007) argue, however, that the loss of male lives in Nepal between 1951 to present due to war and civil conflict is comparatively limited. There have been a series of revolutions and conflicts that have shed a lot of male blood. The "People's War" by the Maoist activists, for instance has caused over 13,000 deaths (1996-2006 civil war). But in relative terms to the overall population these figures are not important enough to severely impact on the sex ratio level.

Following, the three main factors are discussed.

4. NLSS survey based on labour force surveys

During informal discussions with officials from the CBS in October 2018, the author was told that the NLSS structure and sampling method is based on previous labour force surveys. The purpose of these survey is to correctly reflect the labour market and are thus often heavily male biased. Furthermore, the main goal of the NLSS surveys initially was to make judgements on the country's economic (and to some limited extent social) situation. Therefore, the interest was not to perfectly reflect the social and anthropological, but the economic reality of the nation, thus leading to a sex ratio level below the national average. This trade-off was accepted and is probably the main explanation for the disproportionately low sex ratio of the NLSS overall dataset.

5. Physical absence of individuals: absentees/migrants (CBS, 2012; Ministry of Labour and Employment, 2014 and 2018; WB, 2011; CBS, 2011)

There are three main factors as to why the physical absence of individuals significantly lowers the overall sex ratio of Nepal and of the NLSS dataset. First, as previously described (see Section 3.3.2.3.1) a high proportion of the Nepalese population is absent from the HH they originally belong to and 20% to 25% of Nepal's total population is currently residing abroad (CBS, 2012; Ministry of Labour and Employment, 2014 and 2018). More than 84.9% of absentees are in the working age population (15-59 years, NLSS III definition of working population[9]). Figure 3.15 shows the relative propensity for men and women to migrate.

[9] Setting a working population and most economically active age range is difficult. Much literature (e.g. by the OECD or the International Labour Organisation, various economics handbooks e.g. Handbook of Income Distribution) tries to identify the "ideal" age range for what is considered "working population". Also, much literature exists that tries to investigate as to which age range is optimal in defining the most "economically active" part of a population or the age cohort with the highest proportion of employed persons, especially suitable for the analysis of intergenerational mobility.

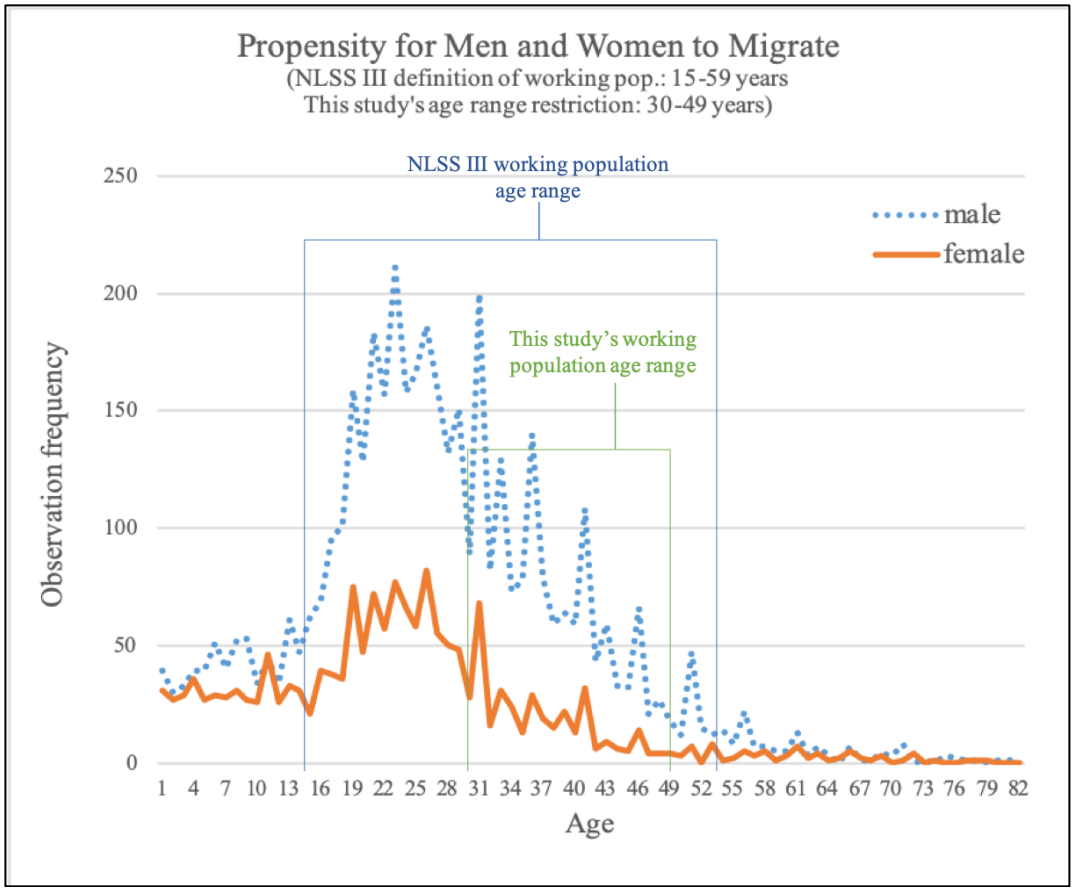
Second, about 32% of absentees fall into the relatively narrow age range of this study's analysed sub-samples (30-49 years of age; author's computation).

Third, male absentees largely outnumber female absentees. More than 70% of all absentees are male and over roughly 32% of total male absentees fall into this analyses' sub-sample group (of the most economically active individually, 30 to 49 years of age). According to the NLSS III working population definition (15 to 59 years), over 75% of absentees are male. When this study's age rage is imposed, over 80% of absentees are male.

So, there is a large proportion of a) absentees, b) who are male, and c) disproportionately do not figure in the NLSS data subset of individuals aged 30 to 49 years. In other words, there is a disproportionate physical absence of the most economically active male population due to international migration. The accumulation of these three absentee factors significantly lowers the sex ratio in my data sub-sample.

Information on absentees is captured in a different section of the NLSS survey and they cannot be included directly in the IOp analysis since they do not have incomplete information on key IOp circumstance variables such as parental education, religion, or ethnicity and caste group belonging.

Figure 3.15: Propensity to Migrate by Gender



Source: Author's computation.

6. Survey response rate: lower propensity of men to report parental education

Out of all the circumstance variables and the various data sample constraints, the propensity to report on parental education is the single most important factor to drive down the sex ratio in this study given the provided data, i.e. male respondents have a lower propensity to report on parental education than female respondents (graphical illustration in Figure 3.16). The "raw" dataset utilised in this

analysis does not include absentees, which decreases the "raw" dataset's sex ratio from the biological norm of 104 to 85 (CBS, 2011), while the 2011 populations census reaches a level of 92. With all IOp data analysis constraints, the sex ratio is driven down to 81 or even as low as 78 depending on the HH composition constraint used.

Furthermore, there is a clear survey attrition problem where sons/daughters and grandchildren consistently do not report on parental education. Despite the interviewer's manual not going into this, due to a time constraint, sometimes surveys do not question information that is already reported on. In this case, if the parents have already indicated their education attainment, then to some extent, this could be directly copied for their sons and daughters. From exchange with CBS officials (personal interviews in fall 2018), this data should have been filled in during the post-data collection phase. However, much of this was not done. Yet, the author filled in missing parental education wherever possible, especially for sons and daughters, and brothers and sisters. Where educational attainment is indicated, this can be copied as parental education for their offspring. So, for example, from 1st (grandparents attained education level) to 2nd generation (HHH/spouse's parental education level), 2nd to 3rd generation, etc. Also, when parental education attainment is indicated for the HH head, then that can be copied to brothers and sisters. More in Section 3.3.2.5.2 on the parental education filling in process.

The above mentioned reasons explain the low sex ratio in the NLSS III dataset, which is further aggravated by the data constraint for the IOp analyses. This undermines the representativeness of the analysis. So, care must be taken when interpreting the results.

The Anderson and Ray (2010 and 2012) papers study "missing women" and the very high sex ratio in India utilising the "skewed sex ratio at birth" (SRB). Their methodology includes a computation of unbiased death rate of women to calculate the rate and absolute estimates of missing women. From there, an "unbiased sex ratio" is calculated and parental education is randomised. These are further used to estimate the absolute and relative number of missing women. In this IOp study, however, this cannot be done for absentees (attrition of un-observables (vs. attrition of observables which can partially be fixed) and a non-random sub-sample), as those emigrants are different to the remaining population. Yet, there are some discussion on attrition in the migration literature, but this goes beyond the scope of this study.

Much is done **trying to improve the sex ratio** bias in the dataset, especially by trying to minimise the dropping of observations due to factors 5 and 6 described above. First, this is done by filling in missing parental information wherever possible. Section 3.3.2.5.2 elaborated with more information on the filling in process.

Second, the previously mentioned FG(2011) data constraints are relaxed in various manners trying to increase the sex ratio and make the analysed sub-sample more nationally representative. While the national level IOP analyses is executed for the strict three data constraint sub-sample (for comparison purpose), two additional population samples are analysed. One that includes all HH members (sex ratio reaches 81) and the other that includes HH heads and spouses, sons and daughters, and brothers and sisters (sex ratio reaches 91). The other levels of IOp analyses in this study, such as for the Development Regions, then use all HH members in their population samples and for the analyses.

Due to male dominance and a large proportion of the most economically active labour force emigration, the descriptive statistics showed that female HH head and spouses remain and they have a higher propensity to report on parental education than their remaining male counterparts. The sex ratio is further lowered when only HH head or spouse are used for the IOp analysis. The descriptive statistics also reveal that the physical absence of sons, sons-in-law and male grandsons disproportionately lowers the sex ratio. And again, daughters, granddaughters and daughters-in-law have a higher

propensity to report on parental education than their male counterparts further aggravating the issue. Therefore, the top priority was to overcome these challenges and fill in parental education wherever possible.

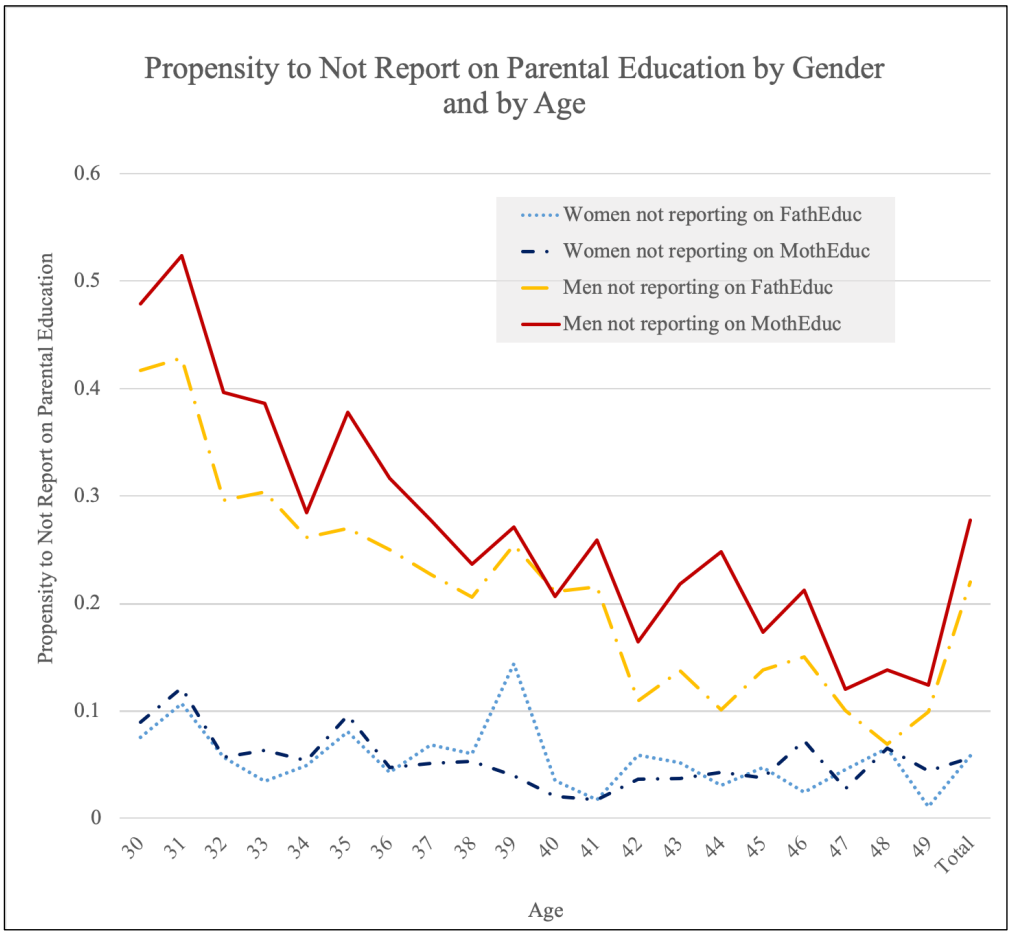
These attrition data imperfections have been acknowledged, investigated into and attempted to overcome to some extent. It is crucial to bear in mind that the sub-sets utilised in the IOp analysis are restricted and biased. Some brief attempts were made as to how these flaws, and especially the two above factors, could impact on the IOp analysis and its results.

Table 3.9: Sex Ratio Comparison Table

Data Source	**Sex Ratio***
"Normal" or "Natural" sex ratio, WDI	106
World average sex ratio, WDI	101
NLSS I (1995/6)	95
NLSS II (2003/4)	92
NLSS III (2010/11)	85
Population Census (2011)	92
WDI (2011) for Nepal	96
My analysis (NLSS, 2011); data sample (original): 1) HH head/spouse 2) 30-49 years old 3) All circumstance variables available	78
My analysis (NLSS, 2011); data sample: 1) HH head/spouse, 2) All circumstance variables available	75
My analysis (NLSS, 2011); data sample: 1) 30-49 years old 2) All circumstance variables available	56
My analysis (NLSS, 2011); data sample (original): 1) HHH/spouse, sons/daughters, brothers/sisters 2) 30-49 years old 3) All circumstance variables available	91
My analysis (NLSS, 2011); data sample (original): 4) All HH members 5) 30-49 years old 6) All circumstance variables available	81

Note: *Number of men per 100 Women

Figure 3.16: Propensity to Not Report on Parental Education by Gender and Age



Source: Author's computation. *Note*: "Don't know" and no replies are combined. The sample is constraint to: age 30-49, parental education Information, for all HH members

3.3.2.4. Dependent Variable Definitions & Aggregation: Per Capita Income and Per Capita Consumption Expenditure

Table 5.3 in the appendix summarises the main empirical IOp literature. It also contains information on which outcome variables are used. As can be seen, a vast majority of studies uses some sort of economic measure (such as income or consumption) to estimate IOp. Mostly the income or consumption measures vary in terms of HH or per capita level, or pre- or post-tax. Few studies use, for instance, education attainment or health condition to estimate inequality in education or health opportunities.

This study uses both per capita income and per capita consumption expenditure as economic outcome, i.e. dependent variables. Chapter 2 of this dissertation elaborated extensively on the literature of the measure of economic welfare, the utilisation of income and consumption as proxies for this and the means of aggregating and interpreting these measures. Section 2.4 focused on the aggregation methodology of consumption expenditure and Section 2.5 focused on the aggregation methodology for income. Both sections also compare the aggregated results and highlight their weaknesses and strengths. Both variables were first computed at the HH level and then at the per capita level. Neither income nor consumption are ideal economic welfare measures and for a complete debate on this refer to Section 2.1. However, using them as separate economic outcome variables in the analyses of this studies is seen as complementary. In sum, given the developing country context of Nepal and the

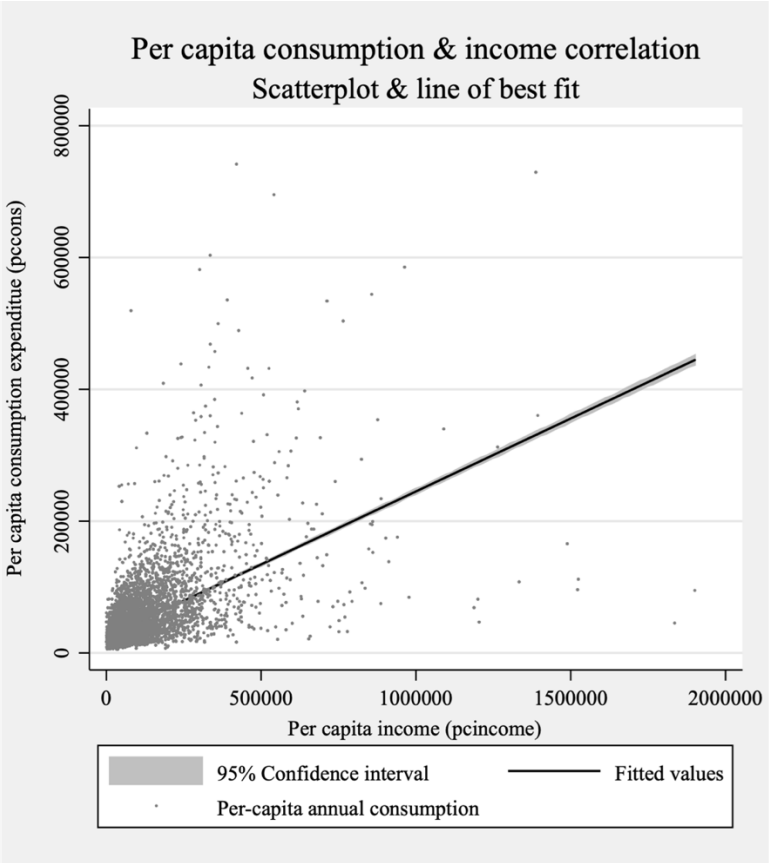
relatively small share of people receiving wage remuneration (less than 50% of income sources, of which less than 67% are in long-term wages), in general, consumption expenditure is preferred over income since it is said to more accurately reflect economic living standards. Yet, both serve as dependent variables in order to check for result reliability and robustness. Thus, two separate panels of results are computed for all levels of the IOp analyses, one using consumption and one using income as dependent variable.

The dependent variables are not solved as simultaneous equations or through other econometric means, but serve as separate dependent variables. In other words, the various IOp analyses are run for consumption and also separately for income. They provide two separate panels of results for each level of analysis. The descriptive and econometric results are then cross-compared in order to check for consistency, due to the different nature of the consumption and income as welfare measures.

There is a strong positive correlation between income and consumption expenditure. The correlation coefficient is 0.48 (significance level: 0.0000). Figure 3.17 shows the scatter plot of *pccons* against *pcincome* with line of best fit (after having dropped values for *pcincome* below 0 and above 2,000,000NRs. for illustrative purpose). This correlation is as expected. As income rises, also consumption rises. At first, consumption rises disproportionately and then, as income rises, consumption does not increase as much anymore.

The dependent variables are tested for statistical difference between them (income (*pcincome*) and consumption (*pccons*)) at the national level including all observations. To test the hypothesis that consumption (M= 45,093NRs., SD=47,589) and income (M=95,750NRs., SD=147,401) were equal, a dependent sample paired t-test was performed. The correlation between the two conditions was examined at $r=0.48$, $p<0.0000$ (positive correlation), suggesting that the dependent sample t-test is appropriate in this case. The null hypothesis of equal consumption and income means was rejected, $t(29,703)=66.42$, $p<0.0000$. The paired t-test rejects the null hypothesis. In other words, the consumption and the income means are statistically different. On average, per capita income is 50,657NRs. higher than per capita consumption.

Figure 3.17: Per Capita Consumption and Per Capita Income Correlation Diagram



Source: Author's computation.

3.3.2.5. Circumstance Variable Definitions & Adaptations

3.3.2.5.1 Circumstance Variable Definitions

In this study, the exogenous **circumstance variables** serve as independent variables. The number of categorical variables under each circumstance variable was kept to a maximum of three. This was done for two main reasons. First, to reduce the number of unobserved types or types with very few observations. Second, to reduce high sample variance, which is important for the non-parametric analysis. According to FG (2011, p. 640), this is crucial, because inter-type conditional means are calculated. Types with few observations have relatively high sample variance and between type estimates are likely to be overestimated, therefore suggesting higher IOp. Dealing with categorical variables involves both quantitative and qualitative thinking. One has to judge carefully the trade-off between the inclusion of a maximum number of types and low type distribution. In order to do so, countless categorical variable combinations were examined both at the descriptive and empirical level. Table 3.10 summarises the circumstance variable definitions, which were finally chosen. The table also presents the expected effects on economic outcome for the RF-OLS regression analyses. Furthermore, at each empirical level of the IOp analyses (national level, Development Region level, urban-rural area level, income and consumption grouping level) the descriptive statistics for the circumstance variables and the type partition distributions are reported.

The literature of IOp emphasises the importance of including as many circumstance variables as possible in order to get the most accurate IOp estimates. However, empirically this is very difficult. Data availability is the biggest obstacle for a complete IOp analyses. Often datasets lack certain

circumstance variables, information or breakdown of them, and lack in sample size. This is why a) different studies use different circumstantial variables, b) their methodologies are forced to be adapt to the nature of their dataset making cross-comparison to other studies difficult, and c) most studies focus on developed countries, because they tend to have the richest datasets. Yet, this study includes seven of the most commonly used circumstance variables and is more complete in that aspect compared to FG (2011), for instance, who only use six circumstances. A brief discussion on each of the used circumstances and their categories follows. Section 3.3.2.5.3 elaborates on the failed attempt to include a key circumstance variable that does not figure in the NLSS III dataset directly, namely parental occupation.

Furthermore, Table 5.3 in the appendix summarises the main empirical IOp literature. It also contains information on circumstances used by different studies. The exact categorization of the circumstances is not included due to dataset divergences, it can, however, be seen that, for instance, parental education, birth region, race or ethnicity and parental occupation are amongst the most commonly used circumstances. Some of the rarer circumstances are "difficulties during childhood", gender of HH head, number of siblings or presence of parents. This is not necessarily due to the lack of data but due to the trade-off between the inclusion of a maximum number of circumstances and low type distribution.

There are three geographical variables in this study: **Development Region**, birth region 1 (urban and rural areas) and birth region 2 (mountain, hill and Terai ecological belt regions). The Development Region variable, however, is not included as a circumstantial variable because it has 5 categories as described in Section 3.3.1 some of which have a relatively small sample population, which would greatly affect type distributing levels and thus biasing the IOp analyses.

Two birth region circumstance variables are used. **Birth Region 1** contrasts urban versus rural area and Birth Region 2 is divided into the three previously described ecological rages of Nepal: mountain, hill and Terai. Both birth region circumstance variable divisions were done as suggested by the CBS (2011, Ch. 1). **Birth Region 2** categories are aggregated using the districts codes applied by the central government as administrative boundaries (CBS, 2011, p. 6). Foreign born individuals receive missing values and are therefore dropped.

The inclusion as circumstances of particularly two variables is heavily debated in the IOp literature. These are **age** and gender. Roemer and Trannoy (2015) argue that both of these variables should figure as circumstance variable if we look at them as exogenous variables one cannot influence (i.e. under the control view). They should also be included under the preference view, because both age and gender are important determinants of preference (i.e. efforts). If we now look at age, Almas et al. 2011) include it as a circumstance in their analysis. However, most IOp literature excludes age due the obvious link between age and preferences. In other words, as we grow up, our preferences change. This study excludes age, despite not including effort in the analysis, because the analyses population sample is constraint to 30 to 49 year olds. The data constraints sections explained the choice of this.

Gender is debated as an exogenous circumstance variable in the IOp literature due to its partial endogenous and exogenous nature. Whether someone is a woman or a man is of course exogenous. However, HH headship is self-reported and is this largely endogenous. Also, despite the interviewer manual's clear definition of what characterises a HH head, the own cognitive perception of the respondents varies largely. Gender can be influenced by the interviewers choice, because mothers for example may have a better overview over HH expenses, occupations, priorities etc. Also, the choice to marry or not is (in most cases) endogenous. For these reason, FG (2011), for example, exclude gender as a circumstance variable in their scalar indices computations and include it in a latter part of their analysis when isolating the individual effects of circumstances on labour earnings. The overall sex ratio in this study (78 for the HH head and spouse data sample) is relatively low compared the

overall NLSS III complete dataset (86; CBS, 2011, p.20) and the Nepal Population and Housing Census (92; 2011). But as previously explained, this is largely due to the applied data sample constraints and the nature of the NLSS III data survey having been based on the labour force survey.

Table 5 gives the descriptive statistics for the circumstance variables for the individuals who satisfy all three sample selection criteria. Over 61% of the observations are women. Furthermore, one of the criteria indicates that analysed individuals must be HH head or spouse. Table 7 indicates that over 54% of the observations are HH heads and 46% are the spouses. For these three reasons and the absence of father's occupation as a key circumstance variable, it was decided to include gender as a circumstance variable into this study's IOp analyses. For the RF-OLS regressions analyses, always a series of regressions both including and excluding gender are run. Depending on the statistical significance of the variable, gender was then either included or excluded from the following partial IOR computation.

In general, however, given Nepal's low development status and caste based societal structure, being a woman is expected to have a negative and being a man is expected have a positive effect on economic outcome and opportunity access.

According to the 2011 Nepal Population and Housing Census (CBS, 2012), Nepal has over 125 reported **castes/ethnic groups**. With over 16.6% (4,398,053), Chhetri is the largest ethnic group, followed by Brahman-Hill (12.2%, 3,226,903). According to Bhattachan (2003, p.14-17) these two groups are also the two dominant caste groups and "the dominant groups indeed control government (executive, legislative and judiciary), politics, security forces, including army, international relations, bureaucracy and mass media. Therefore, as long as there is no revolutionary transformation of society, such dominance would continue in future." In this study, Chhetri and Brahman-Hill are grouped together and coded as the majority/dominating caste and ethnicity categorical variable for the ethnicity and caste circumstance. They make up about 30% of observations at the national level IOp analysis for example. All other ethnic groups can be defined as middle, lower-middle or lower castes. In this study, they are all grouped together and are defined as ethnic minority/dominated caste groups. Further, it is also important to note that even within the Brahman and Chhetri groups (as well as other caste groups), hierarchisation exists and they could be further broken down (Bennett et al., 2008; Levine, 1987; Höfer, 1979). Due to low type distribution, however, the ethnicity/caste circumstance was kept to the above mentioned two categorical variables only.

So, belonging to the main dominating or upper caste group and is expected to have a positive and belonging to a lower or dominated caste groups is expected have a negative effect on economic outcome.

In contrary to FG (2011) this study also uses **religion** as a circumstance variable. According to the National Population and Housing Census (2011), over 84% of Nepalese are Hindu. Therefore, religion was originally divided into two categorical variables, Hindu and non-Hindu. This is also an inspiration from Hassine (2011), who divided Egyptians into Muslims and non-Muslims (similar religious belonging descriptive statistics). However, when analysing the data, religion appeared insignificant and when the religious circumstance variable was divided into three categorical variables (Hindu, Buddhist and Other) significant results were found. Therefore, religion has three categorical variables in this study, Hindus, Buddhists and other religious minorities. Other religious minorities include, for instance, Christians, Kirant, Jain and Shikh.

Buddhists make up over 56% of individuals who are non-Hindu (CBS, 2012) and according to Bhattachan (2003), Buddhists acquire an influential role in Nepalese society and can be viewed as being equal to Hindus (more on this in the various results interpretation chapters). Thus, it is expected that being Buddhist or Hindu has a positive effect on economic outcome, while belonging to a religious minority groups is expected to have a negative effect.

Various **parental education** categorical alternatives were tried out (both descriptive and econometric). The HH survey was conducted in a way that for mother's education, for example, individuals could reply with different degrees of educational attainment (from none, kindergarten to highest university degree), but also with illiterate and literate. The data cleaning process revealed that, for the studied population sample, over 70% of individuals indicated that their parents were illiterate or literate rather than marking their actual educational attainment. Since basic reading and analytical skills are acquired in primary school, over 98% of all individuals fell into no education/illiterate or primary school/literate category. Effort was made to try and find out as to why individuals were given the choice to reply to the parental education question with illiterate, literate or education degree. Yet no satisfying definitive answer was obtained. Furthermore, the descriptive statistics for parental education when divided into three categories indicated that only 8% of fathers and 3% mothers had completed secondary education or above. Apart from the fact that these results are ethically and socially shocking, the categorical type distribution is very low and therefore parental education was divided into only two categorical variables and not into the three categorical variables as originally suggested by EducateNepal (2010) or other empirical research. The two categories are: 1) no schooling, illiterate, and primary education incomplete; and 2) primary education completed and above.

As generally suggested by the literature, the higher the level of education, the more positive its expected effect on economic outcome. Thus, having parents who are illiterate or have not completed primary schooling is expected to have a negative effect on economic outcome and opportunities, while if they have completed primary schooling or continued studying then that is expected to have a positive effect. Some literature suggests that father's education background is more important for the economic welfare and socio-economic mobility of the children than mother's and some literature empirically contradicts this (e.g. Zoch, 2015; Handa, 1994; Chant, 1997; Buvinic and Gupta, 1997).

In general, when circumstance variables have missing values, the observations were dropped. Also, when the answer "don't know (dk)" was given, observations were dropped. In the case for parental education, for example, about 400 observations had "don't know" reported. Since it is not possible to at least identify whether parents were literate or illiterate, and the response is also very different to the "none" response, the observations were dropped.

Table 3.10: Independent Variables (Circumstances)

	Grouping of Circumstance Variables	Expected Effects on y for Reduced Form OLS Regression
Gender		
Category 1	Dominant (Male)	Positive
Category 2	Dominated (Female)	Negative
Ethnicity		
Category 1	Dominant caste/ethnic majority (Chhetri and Brahman-Hill)	Positive
Category 2	Dominated caste/ethnic minority (all others, including middle and lower caste groups)	Negative
Religion		
Category 1	Hindu	Positive
Category 2	Buddhist	Positive
Category 3	Other	Negative
Father's education		
Category 1	No schooling, illiterate, primary incomplete	Negative
Category 2	Primary completed and above	Positive
Mother's education		
Category 1	No schooling, illiterate, primary incomplete	Negative
Category 2	Primary completed and above	Positive
Birth region 1 (U/R)		
Category 1	Urban	Positive
Category 2	Rural	Negative
Birth region 2 (M/H/T)		
Category 1	Mountain	Negative
Category 2	Hill	Positive
Category 3	Terai	Positive

Note: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003, p.17). Religion is divided as by Nepal Consensus Data 2011. Birth region 1, 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).

3.3.2.5.2 Filling in Parental Education

Section 3.3.2.3 on dataset shortcomings already discussed the need to fill in missing parental education where possible. This is first, because the parental education variables is the number one circumstance with the largest amount of observations, which significantly lowers the suitable sample size for the IOp analyses. Second, parental education needs to be filled in in order to improve the sex ratio, because women have a higher propensity to report on parental education than men.

The author was told by CBS officials (informal interview, October 2018) that the parental education sections of the NLSS survey were not asked wherever this information was already given directly or indirectly in other parts of the survey. This information should then have been carried over by the team that cleaned the data. However, the descriptive data analyses showed that this was not done. For the entire dataset, there were over 13,175 missing observations for father's and 17,318 for mother's education.

Missing parental education was filled in wherever possible, especially for sons and daughters, and brothers and sisters. This was done using the education attainment information of each HH member. There are two main stages to this. First, the copying from one HH member to another. Second, analyzing the information given on education attainment.

First, the parental education variable was filled in from using the current education attainment given. Below the 3 analogies for whom parental education could be copied without heavy assumptions:

- 1st → 2nd generation: Father/mother → HH head
- 1st → 2nd generation: Father-in-law/mother-in-law → spouse
- 2nd → 3rd generation: HH head & husband/wife → son/daughter
- Brother/sister HH head parental education → brother/sister

It is important to note that HH membership details are provided around the HH head. So, "father/mother" indicates the HH head's parents. So, for example, from 1st (grandparents attained education level) to 2nd generation (HHH/spouse's parental education level) means that education attainment information that was indicated for the grandparents, was copied as parental education for the HH head. This analogy can only be used for spouses when parents-in-law live in the same HH. Information from HH head and spouse can be copied into sons and daughters' parental education information. The parental education information for HH head could also be copied directly for brothers and sisters of the HH head who live in the same HH. There were some challenges. For instance, some families are polygamous (less than 0.2%), so for them, missing spouse parental education, despite the presence of in-laws was not possible. Furthermore, some HH heads are remarried and have both ex- and current wife living within the HH, etc.

Second, education attainment of HH members is divided into a) completed, b) currently in education and c) never attended school. Information on those who have completed or never attended schooling were added as expected to the parental education variable. There are arguments for and against including the information if someone is currently undergoing education. The proportion of these people in this study's sample with the 30 to 49 years age restriction is about 30%. Their current (even if not yet completed) education status was added to fill in possible missing parental education observation.

There are two categories who disproportionately did not report on parental education. These are sons and daughters, and brothers and sisters (see Table 3.11). The same table indicates that all missing values for these categories were filled in for these two HH membership categories. Furthermore, almost all missing observations were filled in for HH head and spouse.

It is important to note that the filling in process is by nature selective. One knows aperially, in other words beforehand, that the filled in data will be biased and must be viewed at cautiously. There is, however, a trade-off between a very large proportion of missing parental circumstance variables (significantly lowering sample size) and filling in as many of these parental characteristics (which is, however, to some extent selective). Furthermore, the original sub-sample of the NLSS dataset who report on parental education was already self-selected. This is because some people or categories of people have different propensities to report on parental education than others (see sex ratio discrepancy section). Even after filling in the missing information of parental education, the sub-sample is still self-selected and there is a bias. Also, the data is self- reported, because even after filling in, there is a difference between filled in and self-reported parental education. This has to be acknowledged and Table 3.11 highlights the changes made.

Table 3.11: Filling In Parental Education Information

Mother's education									
Relationship to HH head	**No Schooling**	**Primary Educ Incomplete or Above**	**Don't Know**	**Missing or Not Applicable**	**Relationship to HH head**	**No Schooling**	**Primary Educ Incomplete or Above**	**Don't Know**	**Missing or Not Applicable**
RelationshiptoHHhead MothEduc **BEFORE** filling in					RelationshiptoHHhead MothEduc **BEFORE (30-49 years age restriction)**				
head	5,834	578	154	664	head	2,808	276	67	437
husband/wife	4,428	535	112	91	husband/wife	2,357	300	48	70
son/daughter	340	107	5	13,436	son/daughter	73	14	1	490
bros/sisters	262	100	5	196	bros/sisters	57	20	0	47
RelationshiptoHHhead MothEduc **AFTER** filling in					RelationshiptoHHhead MothEduc **AFTER (30-49 years age restriction)**				
head	5,834	578	154	1	head	2,808	276	67	1
husband/wife	4,428	535	112	84	husband/wife	2,357	300	48	61
son/daughter	13,753	130	5	0	son/daughter	559	18	1	0
bros/sisters	458	100	5	0	bros/sisters	104	20	0	0
Father's education									
RelationshiptoHHhead FathEduc **BEFORE** filling in					RelationshiptoHHhead FathEduc **BEFORE (30-49 years age restriction)**				
head	4,386	2,261	285	298	head	2,042	1,202	140	204
husband/wife	3,301	1,663	165	37	husband/wife	1,694	977	82	22
son/daughter	596	2,254	16	11,022	son/daughter	49	67	1	461
bros/sisters	191	252	17	103	bros/sisters	37	53	8	26
RelationshiptoHHhead FathEduc **AFTER** filling in					RelationshiptoHHhead FathEduc **AFTER (30-49 years age restriction)**				
head	4,386	2,261	285	1	head	2,042	1,202	140	1
husband/wife	3,301	1,663	165	24	husband/wife	1,694	977	82	20
son/daughter	11,604	2,268	16	0	son/daughter	509	68	1	0
bros/sisters	294	252	17	0	bros/sisters	63	53	8	0

Source: Author's computations. *Note*: only information for HH head, spouses, children and brothers and sisters are provided. Details for other HH members are available upon request.

3.3.2.5.3 Absence of Parental Occupation and Possible Proxies?

Parental occupation and particularly father's occupation is a key circumstance variable in most IOp studies. However, the NLSS III dataset does not include information on parental occupation. This omission is a limitation to this study and one must bear this in mind when interpreting the IOp econometric results. Furthermore, the absence of parental occupation limits this study's comparability to other empirical studies such as FG (2011), which include father's occupation, but only for 4 out of the 6 analysed Latin American countries. However, significant efforts were made trying to find a suitable proxy that could account for parental occupation.

First, Nepal is a caste based society and most caste groups are profession based, others are of indigenous and geographical nature (Höfer, 1979). The idea was to include parental occupation under the assumption of intergenerational occupation immobility. According to GoNepal.eu "The caste system is divided into four folds which are: Brahman (priests and scholars), Kshatriya (warriors), Vaisya (merchants and traders), and Sudra (labourers). In this system, the membership is both hereditary and permanent. Inter-caste marriage is almost impossible as it carries a social disgrace, especially when it takes place between two castes at the extreme ends of the society." So initially, father's occupation was determined by using caste group belonging into agricultural and non-agricultural sector (Hofer, 1979, Ch. 7 on Occupation and Caste Status). Two assumptions had to be made. A) All individuals work in the occupation prescribed by their caste group belonging, and B) there is complete occupational immobility between generations so that one can infer current caste/occupation group belonging to parental caste/occupation group belonging. These assumptions were judged as too strong, however, for this caste-occupation inference proxy to hold. Especially

during the field work for this dissertation and from exchange with individuals from various public and private institutions, clear indications were given that the caste-occupation system is no longer strict practice in Nepal. This rigorous caste-occupation association, although still present in some rural parts for some isolated caste groups, dates back to over a century ago. These days, particularly urbanisation, trade liberalisation and economic sectoral transformation have heavily impacted on the caste-occupation association and the caste system.

A second proxy attempt was made by assuming that if a HH head or spouse works in the agricultural sector, the father was likely to have also worked in the agricultural sector. Here again, one must accept the strong assumption of intergenerational complete immobility if this proxy is supposed to be used for filling in parental occupation. Furthermore, when individuals work in the non-agricultural sector, no assumption can be made about paternal occupation. So this second attempt to find a paternal occupation proxy was also discredited.

An additional issue with the inclusion of parental occupation on top of the ethnic and caste circumstance is the eminent endogeneity problem between the two of them in the analysis. So overall, father's occupation and possible proxies for it were not included in this study.

3.3.3. Econometric Model Adaptations for the Nepal Dataset

This section highlights some of this study's model and variable adaptations to the NLSS III dataset and Nepal's specific socio-economic context. The author tried, however, to keep this study's IOp analyses as similar as possible to those by FG (2011) in order to allow for comparability. Most changes and adaptations are therefore due to the nature of the NLSS III dataset.

First, one the national level analyses population samples followed the same **data restriction** criteria as FG (2011).

i. Individuals had to be a HH head or spouse (maintained, relaxed)
ii. Individuals had to be aged between 30-49 years (maintained)
iii. Individuals had to have information on all circumstance variables (maintained)

However, two additional population samples are analysed at the national level, which relax i. but keep ii. and iii. constant. The first population sample includes all HH members and the second includes HH heads and spouses, their children and brothers and sisters. This was done to increase sample size. Additionally to the national level analyses, which were also done by FG (2011) for their analysed Latin American countries, this study further disaggregates the IOp analyses by geographic and demographic factors. Since this is Nepal country specific and does not allow for literature comparability directly, the HH head and spouse restriction was also relaxed for these level analyses. All HH members are used in the analyses population samples for these levels of analyses.

Second, the restricted dataset samples were not reweighted. The NLSS III survey was carried out and **weighed** so that the surveyed population was representative of the country. However, data sample constraints are applied across all levels of analyses significantly decreasing the size of the population sample (see Section 3.3.2.2 on the dataset constraints for more details). Therefore, they are no longer representative for the entire country. The observations used in this study's analyses are not reweighed using the provided weight variable, because this would further falsify the results.

Third, there are some **Nepal and NLSS III specific adaptation to the dependent variables**. Chapter 2 elaborates extensively on the convergences and divergences of the aggregation methodology

recommendations for income and consumption expenditure. Much of the particularities in the methodology applied that this study follows is due to the nature of the NLSS data (e.g. measuring units). Some elements also included in consumption, for instance, are Nepal specific, such as the extent of the inclusion of health and education expenditures. Refer to Chapter 2 for more details.

Fourth, the **choice of the circumstance** variables and **their categorisation** is also country specific and much depends on the nature of the dataset. For instance, a few circumstances are added such as caste and ethnic group belonging, and the geographical birth region 2 for the ecological belts mountain, hill and Terai. These are not included by FG (2011), for instance, since they are country specific. Also, the parental education variables were kept to only 2 categorical variables following an intensive investigation into the descriptive statistics and empirical analyses. FG's (2011) parental education variables have 3 categories. Section 3.3.2.5 fully elaborates on the particularities and choice of circumstance variables in this study.

An additional change to FG (2011) was to **include gender** as a circumstance variable. As mentioned before, FG (2011, p. 640) decided to exclude gender from their circumstance variable set due to the HH head's gender endogeneity problem. This problem states that the gender of the HH head is to a certain extent a matter of choice. This may be a matter of choice of the interviewer, but HH headship is also largely endogenous because the individual can decide whether or not to marry. Furthermore, the main woman in a HH may often have a better track of resource allocation than the husband and is therefore more likely to be interviewed. Due to this endogeneity problem, FG (2010) exclude gender as a circumstance variable from their scalar measure analysis. But, because gender is exogenous at an individual level, it was included in a latter part of their analysis where they decomposed the circumstance specific share of opportunity deprivation profiles. In this study, however, gender is included into the analysis as a circumstance variable. There is no need for exclusion since the HH head gender endogeneity problem is limited. Also, this study uses two additional population samples that include also other HH members than just HH head and spouse in which the sex ratio becomes more representative to the national level. Gender is also included in the computation of the scalar IOp indices, some of the RF-OLS estimations and some of the partial IORs. The cost of dropping gender from the IOp indices analysis is larger than in an ordinary OLS estimation, since it is a key independent variable that describes IOp and that is exogenous at the individual level. Another reason why gender is included in this study is that, throughout this study and due to relatively small population samples, there is a large number of unobserved types. The inclusion of gender distributes the population more evenly and balances the many non-observed types. This is important for both the parametric and non-parametric analyses at all levels of the IOp analyses.

As mentioned previously, **father's occupation** is dropped because it does not figure in the survey. Some possible proxies for father's occupation were tried out such as a) occupation sector of the surveyed individual and b) caste group belonging. The assumptions that would have allowed for these two proxies to be suitable were, however, judged as too strong and hypothetical. Thus these proxies for father's occupation were discredited. More details on this in Section 3.3.2.5.3 that investigates into the use of potential parental occupation proxies.

The above mentioned econometric model adaptations to the country specific context of Nepal and the nature of the NLSS III dataset provide both advantages and disadvantages to the analysis of IOp. These adaptations allow a thorough and in-depth analysis of the IOp reality in Nepal. However, they also limit the **comparability of the results** to other empirical studies. In general, the empirical literature on IOp is relatively limited and due to the high demands on data quality and detail, all studies are forced to adapt the basic IOp models to the nature of their datasets. Most modifications are generally in terms of HH composition and circumstance variable inclusion (circumstances and their categorisation). These are the main reasons why the WB's efforts to compute their Human Opportunity Indicators (HOIs) has been restrained to a few countries, and their IOp survey study also retained

serious geographical coverage limits mainly focusing on developed countries, Latin America and some isolated Sub-Saharan African countries (e.g. Barros et al., 2009).

Most empirical IOp analyses are constrained to analysing only **HHHs and spouses** due to the fact that most studies use labour force surveys which are the mainly the ones rich enough to allow for a thorough IOp investigation. Many labour force surveys, however, only capture HHHs and spouses, or only for them capture the necessary circumstances variables. Equally so, most studies focus on computing national level IOp indices (see Appendix 3: IOp Summary of Empirical Literature for more details). That is the reason why the national level analysis in Section 3.4. includes a data sample of HHHs and spouses only. These results are then comparable to other IOp studies to some extent. As mentioned, the HH roaster constraint, however, is relaxed and two additional population samples at the national level are provided and similar results are obtained.

The choice of circumstance variables and their categorisation is different across all IOp studies. Needless to say, this also limits the national level results' comparability, despite also only using HHHs and spouses.

When it comes to the further geographical and demographic disaggregation of the IOp analysis in this study, the comparison to other studies should be done cautiously. Because of the uniqueness in disaggregation the IOp analysis, the author has decided to keep all HH members who satisfy the data restriction criteria in the analyses. Furthermore, the uniqueness of some of the circumstance variables such as caste group belonging and the ecological belt stratification, as well as the circumstances' categorisation, add layers of depth to the analyses, but also limit the results' comparability. Because the country specific contextual adaptations add detail to the understanding of the IOp reality in Nepal, they are important for the subsequent proposition of IOp policies.

In sum, the national level IOp results are to some extent comparable to the empirical literature, however the disaggregated outcomes should be compared with reservation. Given the limited data availability to allow for a thorough IOp analyses across all countries, the vision of having a globally comparable IOp indicator is yet a dream. However, over time, as the coverage and collection of exhaustive data improves one can expect an increase in comparable IOp studies. For now, other basic, much simpler and less complete measures of inequality such as the Gini coefficient will have to suffice when wanting to compare inequality across countries. This does not mean, however, that a thorough country level analysis of IOp cannot serves as reference for formulating egalitarian and socio-economic policies.

This section (3.3) first provided some general background and context information on Nepal, its socio-economic, political and cultural context and geographical conditions. Second, details were provided on the dataset, its imperfections and how they are dealt with, and the dataset constraints that are applied for the IOp study. Then, this section elaborated on the dependent variables, their definitions and particularities (main reference remains Chapter 2). Fourth, an investigation was done into the choice of circumstance variables and how they were adapted for this study. Last, an overview of the main econometric model adaptations for the NLSS III dataset and the Nepalese context was given.

3.4. Inequality of Opportunity Analysis in Nepal: National Level

This section (3.4) focuses on the empirical analyses of IOp at the national level in Nepal. First, this section elaborates on the descriptive statistics at the national level.

Second, the empirical results are given and interpreted. To do so, initially the scalar indices of IOp, then the RF-OLS results and the partial IORs are given discussed. Following the circumstantial specific shares of unequal opportunities for the most opportunity deprived are highlighted and

discussed. The descriptive and empirical findings of the national level analyse are combined and fully discussed.

Finally, various robustness checks are carried out and discussed.

Please note that the national level analysis serves as a reference for all other levels of analyses, Development Region, urban-rural and income and consumption groupings by population quarters. So this section elaborates in detail on all parts of the national level analyses, while the subsequent level analyses sections will only highlight the most interesting aspects of their findings.

3.4.1. Descriptive Analysis: National Level

This section first looks at some general noteworthy descriptive statistics including HH size and composition. It then discusses the type divisions and distributions of the study. Finally, the descriptive statistics for all circumstance variables and their sub-categories are given with an initial interpretation attempt.

The descriptive statistics are important for the subsequent econometric results' interpretations. Also, the descriptive statistics encouraged the **adjustment of the data sample** as well as the analytical methodology. For instance, for the national level analyses, the key reference literature mainly uses HH heads and spouses in their analysis. However, here two additional data samples were added in the analyses: a) the inclusion of all HH members, and b) HH heads and spouses, sons and daughters, brothers and sisters. These are on top of the globally applied data constraints that i) all observations must have complete information on all circumstances and ii) they are aged between 30 to 49 years. These additional data sample analyses sets were done in order to increase the sample size and thus render the empirical analyses more robust and representative.

Another example where the descriptive statistics helped to shape the methodology was for the **division of the categorical variables**. This is the case of parental education. There are many possible ways into which parental education can be sub-divided, many of which were tried out in the empirical analyses. However, type distribution levels are very low the higher the education level. Table 3.12 shows two possible ways in which parental education can be divided. The first one was chosen for all analyses primarily due to low frequency and thus low type distribution. For example, it shows that when parental education is divided into only two categories, namely 1) no education, illiterate and primary education incomplete, and 2) primary education completed and above, the level distribution frequency for category are as low as 39% and 25% for father's and mother's education, respectively. When parental education is divided into 3 categorical variables, (1) no education, unknown and illiterate, 2) primary education incomplete and complete or literate, and 3) secondary education incomplete or more), observation frequency is low for when all HH members are included into the analyses, 23% and 17% for father's and mother's education, respectively. However, when the HH head and spouse constraint is applied, this accounts for only 8% and 3%, respectively. This low level of categorical observation leads to an increase in unobserved types and types with fewer than five observations rendering the subsequent econometric IOp analyses less robust. Parental education has therefore only two categorical variables in the analyses here as well as for all other levels of analyses (e.g. the Development Region or urban-rural area level).

Table 3.7 shows the **household sizes** for all 5,988 households in the dataset. The mean HH size at the national level is 4.79 family members. This is consistent with the Nepal Population and Housing Census (CBS, 2011, p. 3) and the NLSS III Report (CBS, 2011, Vol. 2, p. 41) that estimate mean HH size to be 4.88 and 4.9 individuals, respectively. Figure 5.3 (Appendix 1:) is a diagram showing the evolution of mean HH size since the first NLSS dataset collection in 1995/96. One can see that the

number of smaller HHs has increased since 1995/96 and the number of HH with 5 or more persons has decreased.

Table 3.8 shows the mean household sizes according to different dataset constraints for the IOp analyses. All three samples have two out of the three dataset constraints in common: individuals must have complete information for all circumstance variables and they must be aged between 30 to 49 years. The last dataset constraint applies to the household composition criterion. When all household members are taken into consideration for the IOp analysis on top of the previous two dataset constraints, the mean household size is 5.41. When only HH heads and spouses are taken into consideration, HH size drops to 4.99.

Table 3.14 complements the household size summary tables by showing the descriptive statistics of **HH composition** according to individuals' relationship to the HH head. Over 91% of families in the sample can eb considered as nuclear families. Yet, the practice of joint families is still common. This becomes evident as over 5% of HH members are either sons- or daughters-in-law, 97% of which are female. This is backed by the fact that over 86% of people in the sons and daughters category are male (overall this category makes up about 7.3 % of people in the HHs).

For the data sample that includes all HH members, more than 84% if individuals are either the HH head or spouse. 72% of HH heads are male and 28% are female and this is 6 and 94%, respectively, for spouses. There are a lot less spouses compared to HH heads (36% versus 48% of HH composition proportions. Combining these descriptive stats with the fact that the sex ratio is only at 89.8 suggests that this relatively high number of female HH heads may be due to (seasonal) male labour migration, both domestically and internationally. Identifying HH headship is, however, problematic due to its partially endogenous nature. HH headship can be a respondents personal choice. One may also choose to be single, divorce or marry. The Nepal Population and Housing Census (CBS, 2011, p. 4) indicates that the number of female headed HHs in Nepal has increased by 11% to 26% from 2001 to 2011. Figure 5.2 (Appendix 1:) illustrates the descriptive statistics for HH headship, sex ratio and dependency ratio in Nepal from 1995/96 to 2010/11.

The **mean age** for the HH head and spouse sample is 38.68 years. This appears consistent with the data constraint that all individuals in the sample must be aged between 30 to 49 years.

Table 3.16 shows the descriptive statistics for the **type partition distributions**. As described in the IOp literature review section, individuals are grouped into types where they are homogenous in terms of their circumstances. Overall, type sample partitions vary largely due to variations in sample size. The descriptive statistics suggest that the smaller the sample size and the higher the number of categorical variables the lower the type distribution. See the type distribution section (3.3.2.5.1) for a more elaborate debate on the trade-off between the number of circumstances and type distribution. The national level IOp analyses has 7 circumstantial variables with a total of 16 categorical variables (see Table 3.13 for variable definitions and descriptive statistics). There are over 288 possible type combinations, where individuals are homogenous in terms of their circumstances. For the national level analyses, 167 of these types are observed when all HH members are taken into consideration and the analyses are done for consumption expenditure. The mean type observation is 42 with a minimum number of observations being one and a maximum type observation being 643. As suggested by the literature (e.g. Roemer, 1998, FG, 2011), there is a trade-off between the maximum number of circumstances and categorical variables and low type distribution. Despite having only 7 circumstances and having kept the level of categorical variables to a maximum of three (or two in most cases), there are still some types with low distributions. It is important to note that about 20% of all observed types have fewer than 5, but more than 0 observations. These type distribution statistics do not vary much across all 6 samples. It is difficult to compare these results to the existing literature due to variations in the circumstantial variables used, different dataset constraints and different datasets. For instance, when comparing to FG (2011, p. 642), some of their type observations are higher, some

lower. In general, however, their relative type observations are much higher and often close to 95%, mainly because their sample sizes are much larger. For example, the Brazil PNAD (1996) dataset has over 70,521 observations, all type combinations are observed and only 7% of types had less than 5 observations. This study only includes a maximum of roughly 7,000 observations at the national level.

Low frequency observation for certain categorical variables is a further explanation for the relatively high level of non-observed types and types with fewer than 5 observations. Low distributions and their combination with other categorical variables makes many type combinations unlikely or with extremely low distributions. For instance, the 4 categorical variables with the fewest observations for the HH head and spouse sample, were other religious minorities (7%), Buddhists (9%), urban birth region (10%) and mountain birth region (10%). So, any type that includes one or more of these categorical variables has aperially a smaller type distribution. It is important to note that not just parental education was reduced to only 2 categorical variables, but also ethnicity and caste group belonging due to low numbers of observations.

To sum up, the smaller the population sample and the more categorical variables, the fewer types observed and the higher the proportion of type distributions with less than 5 observations. This is consistent with the literature (FG, 2011, p. 633).

Table 3.13 gives the definitions and expected effects of the circumstantial variables on economic outcome as well as the descriptive statistics for the categorical variables. Only the finally chosen and ideal circumstantial variables with the respective categorical variables are reported. Other categorical combinations are available upon request. The descriptive statistics are given for four different datasets, the last three of which are used for the subsequent IOp analyses.

1) All observations (29,722) with complete information on the economic outcome and circumstantial variables
2) All observations (7,044) with complete information on the economic outcome and circumstantial variables, and aged between 30 to 49 years.
3) All HH heads and spouses (5,899) with complete information on the economic outcome and circumstantial variables, and aged between 30 to 49 years.
4) All HH heads and spouses, sons and daughters, brothers and sisters (6,530) with complete information on the economic outcome and circumstantial variables, and aged between 30 to 49 years.

As previously mentioned, **gender** is a circumstantial variable that needs to be looked at with caution. The nation-wide sex ratio with no data constraints is 89.8, for all HH member with data constraints it is 81.3, for HH heads and spouses it is as low as 78.4 and for the last data sample it is 91.2 (Table 3.15). Data sample 3 has the lowest sex ratio, because it exclusively focuses on HH heads and spouses and they are the most prone to emigrate for economic purposes. At the nation-wide level, more than 85% of individuals are born in the countryside, a large proportion of whom seek economic opportunities in the industrial and urban centres. The female overrepresentation may also be due to the nature of the HH survey as previously discussed. Also, HH surveys traditionally aim at collecting a maximum level of information on HH activities and resources and it is usually the wives, or women, who are the most aware of the these, thus being favoured during the surveying process. Section 3.3.2.3 has a more complete debate on this and other additional reasons as to why women are overrepresented in the NLSS III dataset.

The **ethnicity or caste** circumstance variable was divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Category one represents the most powerful and dominant caste, particularly people from the Chhetri and Brahman-Hill castes. Category two includes people from the lower castes and other ethnic minorities such as non-caste indigenous groups. Across all data samples about 31% of the population belongs to the dominant while 69% belong to the dominated caste groups.

This is consistent with the Nepal Population and Housing Census (CBS, 2011, p. 10), which indicates that over 29% of the population are Chhetri and Brahman-Hill.

The **religion circumstance** variable is divided into three categories and as suggested by the CBS (2011). Across all data samples, more than 84% of the population are Hindu, 9% are Buddhist and 7% are of other religious minority belief. This is consistent with the Nepal Population and Housing Census (CBS, 2011, p. 10), which indicates that over 81% of Nepalese are Hindu, and so on.

The parental education circumstantial variables were divided into two categories that were judge as the most suitable for the subsequent IOp analyses.

1) Parents who have no schooling, are illiterate or have not completed primary education.
2) Parents who have completed primary education, or above.

As described in Section 3.3.2.5.2, missing observations for parental education were filled in by the author for HH heads and spouses, brothers and sisters, and for sons and daughters using the education attainment statistics.

The **parental education** statistics are striking, but not surprising. Mothers' education levels are much lower than those of fathers. It is the lowest for the HH head and spouse data sample, where over 89% of mothers have no education, are illiterate or have not completed primary school, versus "only" 68% for fathers. Fathers appear to have had more access to education and more of them have basic literacy skills (32%). Additionally, only 8% and 3% of fathers and mothers, respectively, have enrolled into secondary education and above (Table 3.12). This is to some extent consistent with the 2011 Population and Housing Census that states Nepal's national literacy rate to be at 66% (male: 75%, women: 57%). It is important to highlight that this study only utilises parental education, while the population census utilises overall literacy and education rate of all individuals and not just for their parents. Therefore, it appears consistent that the literacy level of previous generations was lower than those for the national average today. The population census also illustrates a large gender gap in education access.

The circumstantial **birth region 1 (urban rural) and 2 (Mountain, Hill, Terai)** are divided as suggested by the CBS (2011).

The **birth region 1** categorical variable shows that over 15% of all Nepalese in the dataset were born in urban areas. This is only 10% for the HH head and spouse restricted data sample. These values are lower than those given by the Nepal Population and Housing Census (CBS, 2011, p.3), which indicates that over 17% of people live in urban and 83% in rural areas. Two reasons may explain this divergence. First, this study uses birth and not living region as a reference variable. Second, this study uses the birth region of people aged 30 to 49 years old, who were born before high urbanisation rates and globalisation struck Nepal in the late 1990's after Nepal's "democratisation". Therefore, the descriptive statistics can be said to be consistent. Yet, the rural population share both in this study's data samples and the population census are very high, suggesting a still very low level of urbanisation and urban economic hub development in Nepal.

Across all samples, the **birth region 2** circumstance variable shows that about 10% of observations live in the mountain, 58% in the hill and 32% in the Terai regions. The proportional difference between these areas is mostly consistent with the Nepal Population and Housing Census results (CBS, 2011, p. 3; mountain: 7%, hill: 43%, Terai: 50%). The divergences between this study's data samples and the census may be due to the sample constraints applied in this data set (especially the age criterion) as well as the utilisation of birth rather than living region. The combination of these descriptive statistics then suggest the existence of geographical mobility within Nepal and internationally. People tend to leave the remote rural areas, particularly the mountain region, to live in the urban centres (particularly located in the hill region which hosts the capital and other Special

Economic Zones) as well as the fertile low lands of the Terai region. Also, the Terai region boarders India with which Nepal has an open boarder. Many Nepalese use this region as a last spring board to depart for India, which sees the largest number of Nepalese immigrants compared to all other countries. Exact numbers for India are not available, but are estimated to be about 93% of Nepalese emigrants (Ministry of Labour and Migration, 2014, p.7).

Overall, the descriptive statistics of this study's constrained data samples appear consistent with those of the CBS (2011) and the Nepal Population and Housing Census (CBS, 2011). A slight divergence is found for the parental education and birth region statistics, but brief explanations have been given. Also, the sex ratios diverge, however, as briefly explained in this section and in Section 3.3.2.3 this is largely due to the nature of the NLSS III dataset.

Table 3.12: Parental Categorical Variable Division Comparison, National Level Analysis

		Grouping	**Descriptive stats (%), all HH members** (no data constraints 29,722 obs.)	**Descriptive stats (%), all HH members** (with data constraints 7,044 obs.)*	**Descriptive stats (%), HHH/spouse** (with data constraints 5,899 obs.)**	**Descriptive stats (%), HHH/spouse sons/daughters bros/sis** (with data constraints 6,530 obs.)***
Parental categorical variable division 1	**Father's education**					
	Category 1	No schooling, illiterate, primary incomplete	61.30	65.65	67.60	66.16
	Category 2	Primary completed and above	**38.70**	34.35	32.40	33.84
	Mother's education					
	Category 1	No schooling, illiterate, primary incomplete	75.40	86.29	89.28	86.53
	Category 2	Primary completed and above	**24.60**	13.71	10.72	13.47
Parental categorical variable division 2	**Father's education**					
	Category 1	None/unknown/ illiterate	51.12	61.92	64.72	62.41
	Category 2	Primary incomplete, complete, literate	26.09	26.89	26.92	26.81
	Category 3	Secondary incomplete or more	**22.79**	11.19	**8.36**	10.78
	Mother's education					
	Category 1	None, unknown, illiterate	64.12	83.5	87.47	83.58
	Category 2	Primary incomplete, complete, literate	18.59	10.23	9.05	10.04
	Category 3	Secondary incomplete/or more	**17.29**	6.27	**3.48**	6.37

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. *Note 2*: Parental education was filled using the educational statistics for HHH/spouse, brothers/sisters, and sons/daughters.

*For all 7,077 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

**For all 5,922 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs.

***For all 6,561 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

Table 3.13: Definition and Description of Circumstance Variables, National Level Analysis

	Grouping	Expected Effects on y for RF-OLS Reg.	Descriptive stats (%), all HH members (no data constraints 29,722 obs.)	Descriptive stats (%), all HH members (with data constraints 7,044 obs.)*	Descriptive stats (%), HHH/spouse (with data constraints 5,899 obs.)**	Descriptive stats (%) HHH/spouse sons/daughters bros/sis (with data constraints 6,530 obs.)***
Gender						
Category 1	Dominant (Male)	Positive	47.31	44.84	**43.95**	47.71
Category 2	Dominated (Female)	Negative	52.69	55.16	**56.05**	52.29
Ethnicity						
Category 1	Dominant caste/ethnic majority	Positive	30.43	31.17	32.10	31.43
Category 2	Dominated caste/ethnic minority	Negative	69.57	68.83	67.90	68.57
Religion						
Category 1	Hindu	Positive	83.57	84.32	84.68	84.35
Category 2	Buddhist	Positive	8.70	8.87	8.70	8.95
Category 3	Other	Negative	7.73	6.81	6.62	6.71
Father's education						
Category 1	No schooling, illiterate, primary incomplete	Negative	**61.30**	65.65	67.60	66.16
Category 2	Primary completed and above	Positive	38.70	34.35	32.40	33.84
Mother's education						
Category 1	No schooling, illiterate, primary incomplete	Negative	75.40	**86.29**	**89.28**	**86.53**
Category 2	Primary completed and above	Positive	24.60	13.71	10.72	13.47
Birth region 1 (U/R)						
Category 1	Urban	Positive	**15.40**	12.25	9.93	11.48
Category 2	Rural	Negative	84.60	87.75	90.07	88.52
Birth region 2 (M/H/T)						
Category 1	Mountain	Negative	9.77	9.62	10.40	9.92
Category 2	Hill	Positive	55.33	57.98	58.19	58.04
Category 3	Terai	Positive	34.90	32.40	31.41	32.04

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).
Note 2: Parental education was filled using the educational statistics for HHH/spouse, brothers/sisters, and sons/daughters.
*For all 7,077 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 5,922 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs.
***For all 6,561 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

Table 3.14: Relationship to Household Head, National Level Analyses (Variable *pccons*)

	Descriptive stats (%), all HH members (with data constraints 7,044 obs.)*			**Descriptive stats (%), HHH/spouse** (with data constraints 5,899 obs.)**			**Descriptive stats (%), HHH/spouse sons/daughters bros/sis** (with data constraints 6,530 obs.)***		
Relationship to HH head	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**
Household head	3,383	**48.03**	48.03	3,383	57.35	57.35	3,383	51.81	51.81
Husband/Wife	2,516	35.72	83.75	2,516	42.65	100	2,516	38.53	90.34
Son/Daughter	519	7.37	91.11				519	7.95	98.28
Brother/Sister	112	1.59	92.7				112	1.72	100
Father/Mother	33	0.47	93.17						
Grandchild	4	0.06	93.23						
Nephew/Niece	4	0.06	93.29						
Son/Daughter-In-Law	**374**	**5.31**	98.59						
Brother/Sister-In-Law	64	0.91	99.5						
Father/Mother-In-Law	4	0.06	99.56						
Other Family Relative	22	0.31	99.87						
Servant/Servants Relatives	3	0.04	99.91						
Other Non-Related	6	0.09	100						
Total	**7,044**	**100**		**5,899**	**100**		**6,530**	**100**	

Note: The descriptive statistics are for the variable *pccons*. The table for *pcincome* can be obtained from the author on request.
*For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 5,899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs.
***For all 6,530 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

Table 3.15: Sex Ratio, National Level Analyses

Sex ratio			
Sex ratio, no data constraints	**Sex ratio,** all HH members and data constraints*	**Sex ratio,** HHH/spouse**	**Sex ratio** HHH/spouse sons/daughters bros/sis***
89.79	81.29	**78.41**	91.24

Note: Author's computations using the NLSS III dataset.

Table 3.16: Type Division Description, National Level Analyses

	Per capita consumption expenditure (*pccons*)			Per capita income (*pcincome*)		
	All HH members*	**HHH and spouse****	**HHH/ spouse, sons/ daughters, bros/ sis.*****	**All HH members***	**HHH and spouse****	**HHH/ spouse, sons/ daughters, bros/ sis*****
Maximum number of types	**288**	288	288	288	288	288
Number of types observed	**167**	164	167	170	166	170
Mean number of observations per type	42.18	35.97	39.10	41.80	35.89	38.96
Min. number of observations per type	1	1	1	1	1	1
Max. number of observations per type	643	577	590	702	632	654
Proportion of types with fewer than 5 observations but more than 0 (number of types)	**19.80** (57)	23.61 (68)	20.83 (60)	21.88 (63)	24.65 (71)	22.92 (66)

Note: *For all 7,044 obs. (*pcons*) and 7,106 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 5,899 obs. and 5,958 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs. ***For all 6,530 obs. and 6,623 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

3.4.2. Econometric Analysis for Inequality of Opportunity, Results and Results' Interpretation: National Level

This econometric analysis section consists first of the scalar indices of IOp. Second, it lays out and interprets the results of the RF-OLS regressions of the circumstance variables on economic outcome. It also looks closely at the circumstantial specific shares of unequal opportunities, the partial IORs. Third, the most opportunity deprived types are ranked by their average mean economic outcome and their profiles are interpreted. Last, the section highlights some of the robustness checks that were carried out throughout the analysis.

3.4.2.1. Parametric and Non-parametric Scalar Indices of Inequality of Opportunity

Table 3.17 presents the results for the scalar IOp indices at the national level. Both, the parametric and non-parametric analyses results are reported, where the Absolute Level of Inequality of Opportunity is indicated by IOL (θ_a) and the Ratio of Inequality of Opportunity by IOR (θ_r). The total level of inequality in outcome (the Mean Log Deviation, E_0) is also reported (refer to Chapter 2 for more details on the MLD) and serves as a reference for the interpretation of the IOp indices.

It is advised to look at the results presented in Table 3.17 at three comparative levels.

1) Comparing the **income to the consumption estimates**. Generally, the consumption estimates are preferred, since they are said to better reflect the peoples' welfare (for more details refer to Chapter 2).
2) Comparing the **parametric to the non-parametric estimates**. Both approaches have their advantages and disadvantages (for more details refer to Section 3.2.3).

3) Comparing the results **between the analysed population samples** (for more details on the various restricted data set samples refer to Section 3.3.2.2).

As described in the methodology chapter, the parametric estimates use hypothetical outcome distributions of circumstance variable intra-type effect (mean outcome per circumstance). In order to calculate the non-parametric estimates, the within type distributions were smoothened and separately standardise. Finally, the indices are estimated. Table 3.18 and Table 3.19 report the RF-OLS coefficients that were used to generate the parametric IOp indices.

Total inequality (MLD or E_0) is estimated at 0.24 for per capita consumption expenditure and 0.56 for per capita income for the population sample that includes all HH members. As expected, the income estimate is higher than that of consumption across all 3 population samples. This is largely due to the greater variance in and measurement error of income compared to consumption (refer to Chapter 2 for more details on inequality in outcome estimates). These values are consistent with the FG (2011) estimates that range between 0.351 and 0.462 for their Latin American countries. Also, as expected, the MLD estimate is slightly lower for the data sample that only takes into consideration HH heads and spouses. This is because HH heads and spouses are said to be the age cohort with the highest proportion of employed individuals, thus decreasing variance and improving the level of inequality in outcome. In other words, when all HH members are included, (e.g. children, grandchildren) who still live it home, then that is probably due to economic reasons, increasing the level of inequality in outcome.

It is important to distinguish between IOL (the absolute level of IOp) and IOR (the relative level of IOp to overall inequality in outcome). Across all data samples and for both income and consumption, the IOL estimates are lower than IOR for both the parametric and non-parametric approach. In general, IOL(income) is higher than IOL(consumption) and IOR(income) is lower than IOR(consumption). This is as expected and consistent with the literature (e.g. FG, 2011, p. 643). In fact, sometimes IOR(consumption) can be more than 50-80% higher than IOR(income).

Also, the non-parametric results are always higher than the parametric results. This is consistent with the theory and the literature. Because within type distributions were smoothened and standardised in the parametric approach, as opposed to the non-parametric approach, the variance and thus the levels of IOp estimates are expected to be higher.

Since E_0(income) is higher than E_0(consumption) and IOL(income) is higher than IOL(consumption), when computing the IOp ratios, the IOR(income) is lower than IOR(consumption), possibly leading to confusion and the conclusion that the IOp results are the opposite of those for inequality in outcome when comparing between the income and the consumption panel. The main reason why the IOR(consumption) estimates are higher than the IOR(income) estimates is because of a higher level of within type components of the inequality in income decomposition.

FG (2011, p. 647) state that "inequality of opportunity levels are actually generally lower for consumption than for income (with the exception of Guatemala). IORs are lower for income because the residual inequality in the income distribution is considerably higher, which is consistent with the view that there is greater measurement error, and transitory income variance, in that variable. This suggests the possibility that income-based IORs may underestimate lifetime (or permanent income) inequality of opportunity, since transitory income variance (and likely higher measurement error) is effectively counted as inequality due to "efforts and luck." "

The IOR estimate can be interpreted as the share of inequality of opportunity to the level of total inequality in outcome. In other words, when looking at the data sample which includes all HH members and where consumption serves as the economic outcome variable, between 26% (parametric estimates) and 32% (non-parametric estimates) of overall inequality in consumption expenditure is due to opportunity deprivation. For income, between 10% to 17% of total income inequality is due to

unequal opportunities. As elaborated in the IOp methodology section and by FG (2011), these estimates are said to be lower-bound estimates.

When comparing the indices across the various restricted data samples, one can see that when HH composition is restricted to HH heads and spouses only, the share of opportunity deprivation out of total inequality is lower than for the sample which includes all HH members. This is consistent with the HH headship and spouse arguments previously stated and contributes to the robustness argument of the results.

This study's results are much lower than the results obtained by FG (2011). For instance, their consumption IOR results range from 25% to 54% out of overall consumption inequality across the 6 (or 5 for consumption) Latin American countries they analyse. This is not surprising, because the analysed countries by FG (2011) are more developed and are larger in economic size than Nepal (Brazil ranks the highest at 8, and Guatemala lowest at rank 77 according to WDI, 2011). The Kuznets Curve suggests that as countries develop, they increase inequality (of outcome) until they reach a reflection point. Nepal is one of the 30 least developed countries in the world meaning that overall inequality in outcome variance (and that of opportunity deprivation) is comparatively low to more developed countries. The development of the Kuznets Curve and a possible extension to the concept of IOp can be found in this study's extension ideas section (3.10.2.4).

This argument is further backed by the fact that this study's estimates are similar to those computed by Checchi and Peragine (2010). In their IOp study on Italy, they find that opportunities account for "only" less than 20%. This, however, is no surprise, since Italy is Europe's 4th and the world's 11th largest economy, while Nepal ranks amongst the 30 least economies in the world (WDI, 2011). Italy is a developed country with (relatively) stable political and jurisdictional institutions that obey the rules of checks and balance, while Nepal is an emerging country with an unstable political system and an evolving jurisdictional infrastructure section (3.10.2.4).

It could be suggested, comparing the results of this study to those by FG (2011) and Checchi and Peragine (2010) that the more a country develops the higher the level of IOp until it reaches a certain inflection point. From there on the level of IOp begins to decrease.

Furthermore, one can argue that because Nepal is such an underdeveloped country, the variance of living standards (i.e. consumption and income) across the country at the national level is not very large, and is by far not as extreme as in more developed countries. In other words, all people (or a vast majority of people) are all more or less equally poor. Furthermore, the IOp empirical methodology employed in this study (direct ex-ante approach that follows the path-independence assumption and used MLD as the reference inequality in outcome variable) is relatively more sensitive to inequalities in the lower part of the distribution.

The **statistical differences** between the IOL versus IOR and parametric versus non-parametric estimates were tested. Table 3.91 shows all the results using the paired t-test. In all cases the null can be rejected and the means of the IOp estimates, no matter whether IOL, IOR, parametric or non-parametric, are different from each other.

Furthermore, the statistical differences between the IOL and IOR estimates (parametric and non-parametric, for income and consumption) between all levels of the IOp analyses, i.e. national, Development Region, urban rural area, and by income/consumption population quarter level, were tested. Table 3.95 reports the results of the paired t-test to test for statistical difference between all levels of the IOp analyses. The correlation matrix is available upon request. The coloured boxes within the table show the within level results statistical difference comparison. The comparison of the statistical differences between the national average IOp indices to those of the other three analytical levels is important. The computed national level IOp indices are statistically different from those of the eastern, central and western regions, the urban and rural areas, and mostly with all of the different income and consumption groups by quarter.

To sum up, the consumption estimates for total inequality in outcome are consistently smaller than the income estimates. This is as expected. The non-parametric estimates are consistently larger than the parametric estimates. This is also as expected. The consumption IOR indices indicated that between 26% to 32% of overall inequality in consumption is due to opportunity deprivation. For income, the opportunity deprivation level ranges from 10% to 17% of overall income inequality. Furthermore and as expected, the level of opportunity deprivation is lowest in the HH head and spouse data sample. This is also as expected.

Compared to other studies in terms of 1) parametric versus non-parametric estimates, and 2) income versus consumption estimates, the results are consistent with the literature (e.g. FG, 2011; Checchi and Peragine, 2010). The estimates obtained by the author, however, are lower than those obtained for various Latin American countries obtained by FG (2011) possibly reflecting the extremely low level of development and thus a lower level living standard variations across Nepal. So, from an IOp perspective, at the national level, the share that unequal opportunities contribute to overall inequality in outcome, appears to be not too bad. However, the subsequent sections of this study that disaggregate the national level analyses by various geographical factors and by population grouping, paint a very different picture.

Table 3.17: Scalar Indices of Inequality of Opportunity, National Level Analyses

	Nepal (for all HH members)*	Nepal (HHH/spouse only)**	Nepal (HHH/spouse, sons/daughters, brothers/sisters)***
Per capita consumption expenditure (variable *pccons*)			
Total inequality (E0)	**0.242**	0.244	0.242
Non-parametric estimates			
IOL	0.077	0.067	0.075
IOR	**0.320**	**0.285**	**0.311**
Parametric estimates			
IOL	0.063	0.052	0.060
IOR	**0.263**	0.219	0.250
Per capita income (variable *pcincome*)			
Total inequality (E0)	**0.560**	0.561	0.561
Non-parametric estimates			
IOL	0.094	0.094	0.099
IOR	**0.169**	0.167	0.176
Parametric estimates			
IOL	0.061	0.054	0.061
IOR	**0.109**	0.096	0.109

Note: *For all 7,044 obs. (*pcons*) and 7,065 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 5,899 obs. and 5,920 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs. ***For all 6,530 obs. and 6,583 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

3.4.2.1. Reduced-Form OLS Regressions and Circumstance Specific Shares of Opportunities (Partial IORs)

The RF-OLS regression results of the categorical circumstance variables on per capita consumption expenditure are reported in Table 3.18 and on per capita income in Table 3.19. Countless regression and categorical variable combinations were tested (see Section 3.5.2.4 for robustness checks) and the tables here report the four main regressions that were judged the best for each data sample. The fourth regression (national level 4) excludes **gender** due to the possible endogeneity

problem associated with it. It also excludes Birth Region 2 (Mountain, Hill, Terai) due to the low statistical significance levels and negative correlation between the region (as can be seen in the correlation matrix tables for the independent variables; Table 3.20 and Table 3.21). Gender was singularly excluded from regression 2 and birth region (Mountain, Hill, Terai) was singularly excluded from regression 3. Birth region 2 (Mountain, Hill, Terai) was excluded from regressions 3 and 4.

Overall, no coefficient changes significantly from one regression to another. R-squared also does not change significantly. All coefficients have the expected signs and almost all are statistically significant at the 1% significance level.

The **correlation matrices** (Table 3.20 and Table 3.21) indicate that father's and mother's education tend to positively correlate. Parental education, however, suffers from ethnic caste group belonging and being born in a rural area. Mothers' opportunity deprivation to access education is higher in rural areas than that for fathers'.

Table 3.22 shows the ***partial* IOR** estimates, that is the circumstance specific shares of overall opportunity deprivation (or enhancement) compared to overall inequality in outcome. Table 3.94 in the overall results comparison Section 3.8 summarises the partial IOR results across all four levels of analyses. Section 3.2.3 elaborated on the computation of the partial IORs, but it is important to re-emphasise that these estimates should be referred to cautiously without giving them too much weight since strong assumptions need to be adhered to in order to compute them. The most constringent assumption being that all omitted variables are orthogonal and uncorrelated to the observed circumstances. Also, there are several steps involved in computing the partial IORs meaning that the total of all computed partial IORs do not necessarily add up to 100 percent. Furthermore, the partial IORs are only valid for circumstances. It is not possible to make a precise judgement on how much of the remaining outcome inequality is due to effort or other factors such as inherited skills, talents or luck. Nevertheless, the partial IORs do give some general indication of the contributing amplitude of circumstances on the level of opportunity deprivation or enhancement and are complementary to the scalar IOp estimates and the RF-OLS regression results.

The correct interpretation for partial IORs is important. It is the maximum estimated share of a categorical circumstance out of overall inequality, giving an indication to the extent of which the circumstance is an important element to boosting or hindering opportunity access. If the RF-OLS coefficient estimate is negative, the partial IOR can be interpreted as the share of aggravation of opportunity deprivation. When the categorical circumstance seen as enhancing opportunities (i.e. positive RF-OLS coefficient), then the partial IOR is the highest estimated share of the circumstance to facilitate access to and boost opportunities.

Overall, the partial IORs for consumption are lower than for income, which is as expected. Mainly the partial IORs for consumption are referred to. At the national level and for the consumption panel, ethnic and religious minority belonging, parental education and being born in a rural area account for up to 76% of overall consumption inequality and are suggested to be important explanatory factors for opportunity deprivation. The later analyses will show that the national level partial IORs are relatively low and homogenous compared to when they are disaggregated by development region or urban rural area.

As expected, belonging to a dominated **caste group** appears to impact on consumption and income levels at the 1% statistical significance level. This is slightly more so for income than consumption, which goes in line with the previously described literature on welfare measurement. The ethnic minority partial IOR indicates that out of overall consumption inequality, more than 14% is due to ethnic minority or discriminated caste group belonging.

Bhattachan (2003) emphasise that still today, individuals belonging to ethnic minorities and particularly lower, lower middle and middle caste groups, still suffer from discrimination. Despite Nepal's relatively progressive legal framework to enhance egalitarian human rights, caste discrimination is still an important issue in Nepal today. There are various reasons as to why caste

group belonging and consequent discrimination and favouritism is still enrooted in Nepalese society today. First, Nepal's limited level of economic development and relatively large population of over 26 million means that people are in fierce fight for economic positions. It is in the interest of the most economically and politically influential people, who historically belong to the upper castes, to maintain their economic status for themselves and their offspring. The level of leverage, and lobbying from the outside and within the yet fragile political institutions is incredibly strong.

Second, the government has limited resources to execute the progressive laws and regulations against caste and any other sort of (e.g. gender) discrimination.

Third, caste group belonging does not only give economic but also social status benefits, and behaviours take generations to change. These outweigh the economic benefits particularly in the rural areas and the social constraints are within communities. Ethnic and dominated caste groups face social stigmatisation on a day-to-day basis, since they are seen as inferior or "impure" by society and possible contact (even through purchasing merchandise) from a person of a superior caste is said to render them impure (Höfer 1979). A Nepalese saying says that for a Dalit (someone from the untouchable caste, i.e. the most inferior) to sell fish, it must still be alive. But for fishermen from the middle or upper caste to be fisherman, they can simply sell the fish meat of a dead fish. This is because, if dead meat is purchased from an untouchable by someone of a superior caste, it is seen as a bad omen. Another popular saying goes "a father can never become an enemy, but a Newar (a caste group particularly in the Kathmandu valley) can never become a friend". These social stigmatisations of course find their root of the top interest groups to maintain in power. Even within the upper caste, there is great hierarchy and one can find Brahman or Chhetri fishermen, who, however, have an interest to hinder their competition for selling fish. Thus allowing Chhetri to sell just the fish meat, while the Dalits need to sell the fish still alive and fresh.

Fourth, changing the legislation is important, but changing the mindset and behaviours of people is more difficult. So, for instance, in public schools people from all caste groups are mixed. While the children do not discriminate much amongst them based on caste belonging, the teachers are more prone to do so (interview with Cheer Japan Oversees Cooperation Volunteers, Pokhara office, October 2018).

Fifth, social networks act as socio-economic security net. Nepal is still an underdeveloped country with still unstable and fragile political institutions that are still at an infant stage of democratic development. The Nepalese government has limited resources at its disposal and Nepal cannot yet be classified as a social welfare state, like Germany or Switzerland, for instance. Since governmental institutions cannot guarantee social security, individuals have to rely on other social safety networks. These networks tend to be within the groups of people that they belong to, upfront the family, closely followed by their caste group belonging or even the religious group (amongst others). Ezemenari and Joshi (2019, p. 6) state that "personal networks are an essential element for youth and adults who seek employment and the quality of these networks are often linked to levels of income and social status". So, unless the government can begin to guarantee a certain degree of socio-economic security, people are heavily dependent on their social group belongings for economic and social security and will continue to favour individuals of people from within their group, discriminating individuals from outside their group.

Sixth, the caste group discrimination, like any other form of discrimination, is particularly due to low levels of education. Unless the overall education level for everyone across the country upgrades significantly from the currently very low levels, within and between community discrimination will prevail.

Last, the nature of the data sample is biased for individuals aged 30 to 49 years. They belong to the older generations before Nepal and international aid began to focus heavily on social inclusion and education projects across the country in the early 2000s. So, the more educated and "new mindset" people are not yet covered by this study's population sample of the 2010/11 NLSS dataset. Once the

NLSS IV dataset becomes available (most likely in 2020), it would be interesting to redo this analysis and compare the results' evolution.

Interestingly, when Birth Region 2 (Mountain, Hill, Terai) is excluded from the RF-OLS regression, the coefficient increases its negative impact on economic outcome. When referring to the correlation matrices tables, it shows that Birth Region 2 (Mountain , Hill, Terai) correlate with the dominated caste variable. There is a positive correlation with being born in the Terai region. This is plausible since there most people work and can make a relative living. Yet the correlation with the Hill region is negative. This is also plausible, because low caste individuals in the Hill region tend to rely more on generating revenue outside of the primary sector, and as the IOp analyses of the later sections show, it is in the urban centres where the most disadvantaged face the highest level of discrimination.

The two categorical **religious circumstance variables** are statistically significant at the 1% level for both economic outcome panels, that is income and consumption. They have the expected signs and remain relatively constant across the four regressions in each economic outcome panel. The amplitude of the coefficients are higher for the income than for the consumption panel. On the one hand, the Buddhist coefficient for the national level 1 regression for income is about double that for consumption and it is positive (0.286 to 0.136, respectively). On the other hand, belonging to a religious minority group other than Buddhism appears to have a negative impact on economic outcome and the coefficient is more than double for when income rather than consumption is the economic outcome variable (-0.232 and -0.103, respectively). The fact that the income coefficients are higher than the consumption ones goes in line with the literature on the measure of economic welfare. Furthermore, the religious minority other than Buddhists partial IOR indicates that out of overall opportunity deprivation more than 17% is due to ethnic minority or discriminated caste group belonging. It is the highest partial IOR of all circumstances and it is consistent across all three national population samples. Being Buddhist is seen to contribute 6% to inequality, thus enhancing opportunities according to the partial IOR estimates.

There are a few reasons as to why being Buddhist is seen as a positively impacting circumstance on economic outcome. First, Bhattachan (2003) describes that Hindus are the dominant religious group in Nepal and that Buddhists hold a highly influential, privileged and outspoken role in Nepalese society, particularly compared to all other religious minority groups. Bhattachan (2003, p. 32) states that "Hinduism is the state religion and Hindu's view that Buddhism is a part of Hinduism". So, Buddhists in general are respected and afford privileged status in Nepalese society.

Second, the Nepalese caste system is enrooted in Hinduism. Buddhism, however, sees all living things as equal and does not coincide with the traditional classifications of the Hindu caste system. This is why some people, especially from the Dalit community (the untouchables), a) convert to Buddhism or b) send their children to Buddhist monasteries. Converting to Buddhism allows them to "lay off" their caste group belonging, particularly when they move to more anonymous urban centres to do so. Yet, for administrative purpose, all Nepalese must have a caste group belonging, and so also do Buddhists. Should they have converted (also from a young age), then they conserve their previous caste belonging on paper. Should they have been Buddhist for generations, there are some Buddhist "caste groups" that do not fall into the roaster of the strict hierarchical Hindu caste grouping and are treated as apart (Interview with Thubten Jikdol, Lead Monk at Kopan Monastery, Kathmandu, 7 October 2018). When parents from the untouchables send their children to Buddhist monasteries it allows them to access a relatively higher level of education and escape absolute poverty. The monetary schools are known to provide a relatively high level of education, especially compared to the public school that the Dalit children would usually attend.

Third, the Buddhist community in Nepal is very well organised for historical reasons, in part due to the large Tibetan community in Nepal. The Tibetan government in exile (officially referred to as the Central Tibetan Administration) in India started to institutionalise educational training for Tibetan refugees world-wide since the very beginning of its founding in 1959. This has guaranteed a

relatively good access to relatively high level of education for Tibetan refugees also in Nepal. While the Tibetan refugees in Nepal are virtually invisible from an administrative standpoint since most of them are not even granted refugee status and are thus denied all legal status, they have been hugely influential for the Buddhist communities and teachings around the country. This in turn has allowed Buddhist schools and high level scholars to seize a relatively influential role amongst the Nepalese elite. Also the well doings of the Buddhist monasteries have elevated the good reputation of Buddhists across Nepalese society for almost 60 years. (nutshell summary of elaborate interviews with various lead monks at various Buddhist monasteries in Nepal, as well as Tibetan refugee camp representatives of Pokhara, Nepal; October 2018). Furthermore, Lumbini in the south of Nepal is said to be the birthplace if Buddha, which in turn elevates the prestige and reputation of the most senior Nepalese Buddhist preachers amongst Buddhist followers internationally and thus also enhances their recognition in Nepalese society and affords them an elite status.

Fourth, the common practice of joint families is deeply enrooted in Hinduism and traditional Nepalese culture. This is less so for Buddhist families. Even when the economic situation of a Hindu family allows them to be nuclear, many family members will still de facto live with them and it is expected that the wives take care of everyone and thus "only" engage in unproductive labour. Less frequently in Hindu but more frequently in Buddhist nuclear HHs, wives seek work and actively engage in the labour market. These HHs tend to be proportionately better educated and practice more gender equality.

Interestingly, the value of the Buddhist coefficient rises as birth region (Mountain, Hill, Terai) is excluded from the OLS regression. A possible reason is that virtually all Buddhists in Nepal live in the Hill and Mountain ranges. Therefore there could have been endogeneity between Buddhism and these birth regions. However, when looking at the correlation matrices between the circumstance variables, all birth region variables seem to be uncorrelated to the religious circumstance variables.

Belonging to a minority religious group other than Buddhism appears to have a negative impact economic outcome. This is as expected. Similarly to the caste discrimination interpretation paragraph, the arguments of the lack of social welfare state and the low levels of general education, hold for religious minority discrimination.

As expected, the higher the level of **parental education** the more positively this impacts on economic outcome. Both parental education coefficients, mother's and father's education, are statistically significant at the 1% level for all regressions across all data samples. Their amplitudes do not change greatly from one regression and from one population sample to another, and they are statistically different from each other (paired t-test).

Interestingly, when consumption expenditure is the dependent variable father's education appears more important than mother's education. For income, however, mother's education appears to be more important. When looking at the partial IORs, parental education combined explains more than 30% of overall consumption inequality. Contrary to the RF-OLS results, the partial IORs suggest that mother's education appears to impact marginally more on opportunities than father's. For both economic outcome panels for the RF-OLS regressions and the partial IORs, the difference between mother's and father's education are statistically different.

In other words, the difference in means between these two variables is statistically different from zero. This allows us to conclude that in the consumption panel, father's education appears to more positively affect children's future consumption expenditure than mother's education. This result suggests when it comes to minimum living standards (i.e. the actual consumption level) of HHs, father's education prevails for children. This, however, goes against much of the academic literature that suggests that mother's education levels are more important for children than father's (e.g. Zoch, 2015; Handa, 1994; Chant, 1997; Buvinic and Gupta, 1997). There are various explanations as to why father's education appears to be more important than mother's education in the Nepalese context. Most of these were discussed during the fieldtrip of the author to Nepal with individuals from various public

and private institutions in Nepal. First, despite a vast majority of Nepalese men out-migrating to seek work in the urban centres or abroad, which would suggest female empowerment, de-facto they are still the decision makers in the family. This is because 21st century technology guarantees a regular and often daily contact between the remaining wives and the husbands. So, men still get to take most of the main decisions, that is resource demanding decisions. Second, male outmigration would suggest female empowerment at home. However, women appear to be triple burdened by this. Now they need to a) take care of the children, b) take care of the parents in law and other relatives due to the practice of joint families, and c) most importantly are responsible for generating economic revenue for the families survival because the frequency and the volume of remittances are mostly uncertain. So, the children are mostly taken care for by the grand-parents, who tend to be more conservative and traditionally mindset. These family pressures are often more aligned towards those of the traditionally more conservative fathers than mothers. Third, on important consumption expenditure matters that imply a much higher financial burden and opportunity costs, such as higher education expenses (especially when in urban centres or abroad), father's still take decision. Mothers tend to solely be able to decide on smaller consumption expenditure, such as primary schooling or food related expenditures.

The fact that father's educations appears more important than mother's must, however, be regarded with caution. First, because, over 35% of fathers while only 14% of mothers have at least some basic education. In fact, gender equality and particularly women's education enhancement programs only began to gain serious momentum in Nepal in the early 2000's, that is after the international community began immensely funding such programs (interview, Professor Sharma, Kathmandu University, 8 October 2018). So the effect of these gender educational programs do not show yet in the NLSS III dataset and not in this study's analyses sample population which restrains age to 30 to 49 year olds. Time will show if these programs are effective. Second, because women are over represented in the sample (55%). (refer to the descriptive statistics section for more details).

For the income panel, however, mother's education attainment appears to be more important than father's education. FG (2011, p. 645) find similar results for all the 6 Latin American countries they analysed. The results have to be compared cautiously, however, since FG (2011) use an additional key circumstance variable (father's occupation) that this study does not include. Also, this study has a Nepal specific additional birth region (Mountain, Hill, Terai) and a religious variable, which FG (2011) do not use. Furthermore, FG (2011) utilise three categorical variables for the parental education circumstance. In fact, much empirical literature exists that suggests mothers education to more affect child future economic output than fathers education (e.g. FG, 2011; Handa, 1994; Chant, 1997; Buvinic and Gupta, 1997).

Mother's and father's education appear to correlate (refer to the correlation matrices). Yet, these two variables are included into the regression and IOp analyses since they are considered as key circumstantial variables in the literature. The rank correlation coefficients between them are statistically significant.

Birth region 1 (urban rural area) appears to negatively impact on economic outcome, more so on income than consumption. The results are statistically significant at the 1% level and are consistent across all sample populations. This result is as expected, because over 65% of Nepal's GDP comes from industrial activity and the service sector (WDI, 2011). Most of these are located in urban areas, particularly in and around the country's capital Kathmandu in the central region. Also, being born in a rural region is expected to negatively impact on economic outcome, because Nepal's level of development is extremely low. The provision and quality of hard and soft infrastructure drastically deteriorates as one leaves the urban centres and the further one goes away from the capital and the central region. Furthermore, the circumstance specific partial IOR result suggest that being born in a rural area impacts on opportunities by over 14 percent of overall consumption inequality. Section 3.6 further investigates into the urban rural area effect on economic outcome and IOp levels.

The **birth region 2 (Mountain, Hill, Terai)** categorical variables are not as significant as the ones for birth region 1 (urban, rural). The corresponding correlation matrices indicate that hill and Terai are the only two variables that appear to be correlated (negatively). This is consistent, since birth region is mutually exclusive. Therefore, the birth region 2 circumstances were excluded from regressions 3 and 4 for each analysed sample.

Whether someone is born in a particular ecological belt does, however, appear to affect ones future economic potentials. Terai birth region is not statistically significant or only at the 10% level. However, being born in the hill belt appears to positively affect economic outcome at a 1% statistical significance level and that more so for income than for consumption. Furthermore, the circumstance specific partial IOR results suggest that being born in the hill belt positively impacts on opportunities by accounting for over 13 percent of overall inequality. The RF-OLS and the partial IOR results of these categorical variables also imply that being born in the mountain belt has a strong negative impact on economic outcome.

The three ecological belts of Nepal divide the country from east to west based on topographical and ecological, and not economic or social characteristics. Even though the hill range does not have economic power in the economic sectors (e.g. agricultural), it contains the country's capital, Kathmandu, and other urban centres, which host the country's prime Development Regions (i.e. with the highest levels hard and soft infrastructure). These areas are also hubs for Nepal's economic secondary and tertiary sectors (NLSS III Reports, Vol. 1 and 2; CBS, 2011). This mainly explains why being born in the hill belt positively effects economic outcome. It is important to bear in mind, however, that the agricultural sector contributes over 35% of Nepal's total GDP (WDI, 2011), most of which is located in the fertile lowlands of the Terai belt. This partially explains as to why being born in the Terai belt, only somewhat "neutrally" impacts on economic outcome, compared to someone being born in the hill (strongly positive) or the mountain (strongly negative) belts. The economic sectorisation of the hill and Terai belt also explains why the coefficients for the hill belt are higher for income than for consumption. This is because the economic activities there are more prone to monetary reward while the traditional agricultural sector often still pays in-kind and does not allow for extreme consumption level variations.

For the consumption panel, the coefficient for being **female** is statistically significant at the 1% level only for the HH head and spouse, and for the third sample. For the sample that includes all HH members, it is only statistically significant at the 5% level. For the income panel, all gender coefficient estimates are statistically insignificant. Looking at the partial IORs of gender, the results suggest that when gender is statistically significant it can explain over 15 percent of overall income inequality. This suggests that gender, however, matters more for consumption than for income to some extent. This goes in line with the literature that suggests that the role of women is more important for intra-household resource allocation (e.g. Handa, 1994; Chant, 1997; Buvinic and Gupta, 1997). In fact, the national level IOp analyses has the most significant (consumption) panel gender coefficients. When the IOp analyses are, however, disaggregated in the following sections, the gender significance completely disappears.

There are several possible explanations for the relative insignificance of the gender circumstance variable on income (at the national level and for the subsequent other level analyses). First, as previously mentioned, the endogeneity problem associated with gender and the self-selection process of HH headship, etc. Second, the overrepresentation of women in the NLSS III dataset, since the survey is based on the structure of labour force surveys which tend to focus more on the work force, most of whom are male rather than trying to be nationally representative for social and anthropological factors. Third, the disproportional overrepresentation of women in the sample for the IOp analyses. So, where originally a man was the HH head and possibly dominated the wife, due to (seasonal) male emigration the wife becomes (at least temporarily) the HH head and is therefore less exposed to gender discrimination. This means that the possibly significant gender circumstance variable becomes

insignificant. Fourth, women are often defined as unproductive labour, e.g. by doing housework or looking after the children, therefore not directly contributing to the level of economic welfare. Nepal, being an underdeveloped country and where most HHs engage in some agricultural work, most women assist without remuneration on domestic food production. All these factors are partial explanations as to why the female categorical coefficient is so low and with such low statistical significance in the income panel.

While doing fieldwork in Nepal, the author was struck by the efforts that are committed to gender equality at all levels of society, such as at the educational, labour force or domestic empowerment level. From the numerous interviews with individuals from local and international IOs, INGOs, public and private institutions, it became apparent that significant progress has been made in terms of gender equality over the past two decades (that is 5 to 10 years before the collection of the NLSS III survey). This is largely due to male emigration, which willingly or unwillingly empowers women to take more HH decisions more actively and consciously. Women begin to play a more important role in society at an economic, but also a societal level. For instance, the number of divorces filed by women is continuously increasing. Yet, this study's sample is constrained to women aged between 30 to 49 years old, and the most recent gender equality societal developments only stated to gain momentum in the early 2000's and particularly for the younger generations.

The **R-squared** value of this study is around 0.2 for all consumption regressions and over 0.1 for the income ones. This means that the chosen variables explain about 20% of the consumption expenditure and only 10% for the income values. While it may be argued that these R-squared values are very low, compared to the literature they are relatively similar (e.g. FG, 2011). These low R-squared values are largely due to a) a relatively low sample size, and b) a large number of unobserved variables. Probably the largest proportion of unobserved variable is the effort variable, which was dropped in this study's IOp analysis due to the endogeneity problem between circumstances and efforts. Refer to Section 3.1 for more details on this and other reasons as to why effort is dropped and implicitly included into the error term. Furthermore and as already explained, father's occupation was judged from the beginning as being crucial in IOp analysis, but was not included in the analysis due to the nature of the data. It is likely that there are other omitted circumstance variables such as mother's occupation, number of siblings, or public goods accessibility. For example, in their calculation for the WB's Human Opportunity Index, Molinas et al. (2010) suggest the inclusion for number of siblings and the presence of both parents as circumstance variables.

The **regression equation** for regression 1 in the consumption panel and the HH head and spouse sample, which contains only statistically significant coefficient, is as follows:

$$conso = 10.682 - 0.208\ ethmin + 0.128\ buddhist - 0.112\ otherrel + 0.361\ fatheduc + 0.241\ motheduc + 0.083\text{hillBR} - 0.428\ rural + 0.045 female$$

where:

conso	Per capita consumption expenditure
ethmin	Ethnic minority/ dominated caste group
buddhist	Buddhist
otherrel	Religion other (minority)
fatheduc	Father primary incomplete or above
motheduc	Mother primary incomplete or above
hillBR	Hill birth region
female	Female
rural	Rural birth region

Overall, when analysing the RF-OLS coefficients and combining them to the extent possible with the partial IOR estimates of circumstance specific shares of opportunities out of overall outcome inequality, one can conclude that being born in a rural region, having illiterate parents, belonging to a minority religious group other than Buddhism and being a member of an ethnic minority will have a significant negative impact on individuals' economic outcome. Parental education accounts to over 30%, ethnic or caste discrimination to over 14%, belonging to a religious minority to over 17% and being born in a rural area to over 14% of overall opportunities. Together, these four circumstances alone explain over 76% of unequal opportunities. Furthermore, the national level analyses are the only ones where gender is statistically significant. Being a woman is expected to have a negative impact on economic outcome and when referring to the partial IOR results gender appears to explain more than 15% of overall consumption inequality, either in favour for men, or to the disadvantage of women. All RF-OLS estimations have the expected signs.

When combining the regression estimate interpretations with the IOL and IOR estimates from the previous section (between 24% to 29% of overall consumption inequality, and 12% to 15% of overall income inequality is associated with unequal opportunities), one can argue that an individual who accumulated all the above mentioned negatively impacting circumstances would significantly suffer from opportunity deprivation. These exogenous factors would predict one's economic wellbeing by up to almost 30% on average at the national level. These above elaborated results and interpretations are frustrating. Compared to the literature, however, the IOp results are not that daunting. Yet, the subsequent sections disaggregate the national level of IOp and paint a much more dramatic picture of a heterogenous Nepalese society in terms of population groupings and geographical divergence.

Furthermore, when combining the descriptive statistics with the econometric IOp indices, the partial IORs and the RF-OLS regression results, one can assume that geographical mobility can be seen as a partial equaliser to IOp. More on this later.

Table 3.18: Reduced-Form OLS Regressions of Circumstances on Per Capita Consumption Expenditure, National Level

Circumstances	**for all HH members**				**for HHH/spouse only**				**for HHH/spouse , brothers/sisters, sons/daughters**			
	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)	ln(pccons)
Ethnic Minority / Dominated Caste	-0.188***	-0.189***	-0.223***	-0.225***	-0.208***	-0.209***	-0.232***	-0.235***	-0.200***	-0.202***	-0.230***	-0.233***
	(0.019)	(0.019)	(0.018)	(0.018)	(0.021)	(0.021)	(0.020)	(0.020)	-(0.020)	-(0.020)	-(0.019)	-(0.019)
Buddhist	0.136***	0.135***	0.167***	0.166***	0.128***	0.126***	0.145***	0.144***	0.135***	0.133***	0.159***	0.158***
	(0.030)	(0.030)	(0.029)	(0.029)	(0.033)	(0.033)	(0.032)	(0.032)	-(0.031)	-(0.031)	-(0.030)	-(0.030)
Other Religion	-0.103***	-0.103***	-0.100***	-0.100***	-0.112***	-0.113***	-0.111***	-0.111***	-0.100**	-0.100***	-0.098**	-0.098**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.040)	(0.039)	(0.039)	(0.039)	-(0.039)	-(0.039)	-(0.038)	-(0.038)
Father's Education: Primary Complete or Above	0.347***	0.348***	0.351***	**0.352*** **	0.361***	0.361***	0.362***	0.363***	0.362***	0.363***	0.365***	0.366***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)	(0.021)	-(0.019)	-(0.019)	-(0.019)	-(0.019)
Mother's Education: Primary Complete or Above	0.269***	0.263***	0.263***	**0.255*** **	0.241***	0.236***	0.234***	0.228***	0.264***	0.254***	0.257***	0.245***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.033)	(0.033)	(0.034)	(0.034)	-(0.027)	-(0.027)	-(0.027)	-(0.027)
Rural Birth Region	-0.483***	-0.483***	-0.494***	-0.494***	-0.428***	-0.426***	-0.430***	-0.428***	-0.467***	-0.464***	-0.476***	-0.473***
	(0.027)	(0.028)	(0.028)	(0.028)	(0.034)	(0.035)	(0.035)	(0.035)	-(0.030)	-(0.030)	-(0.031)	-(0.031)
Hill Birth Region	0.103***	0.104***			0.083***	0.084***			0.103***	0.104***		
	(0.029)	(0.029)			(0.030)	(0.030)			-(0.029)	-(0.029)		
Terai Birth Region	-0.008	-0.009			0.000	-0.002			0.004	0.000		
	(0.030)	(0.030)			(0.031)	(0.031)			-(0.031)	-(0.031)		
Female	**0.032** **		**0.038** **		**0.045*** **		**0.049*** **		0.049***		0.056***	
	(0.015)		(0.015)		(0.017)		(0.017)		-(0.016)		-(0.016)	
Constant	10.709***	10.728***	10.787***	10.810***	10.682***	10.707***	10.741***	10.769***	10.694***	10.718***	10.771***	10.800***
	(0.043)	(0.042)	(0.035)	(0.033)	(0.049)	(0.048)	(0.041)	(0.040)	-(0.045)	-(0.044)	-(0.037)	-(0.036)
Observations	7044	7044	7044	7044	5899	5899	5899	5899	6530	6530	6530	6530
R-squared	**0.216**	0.216	0.21	0.209	**0.175**	0.174	0.171	0.17	0.204	0.203	0.199	0.197
R-squared (adjusted)	0.215	0.215	0.209	0.209	0.174	0.172	0.17	0.169	0.203	0.202	0.198	0.196
F stat	202.775	227.853	239.362	278.641	131.084	147.287	157.306	183.19	176.661	198.705	210.771	245.49

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 7,044 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses.

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 5,899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses.

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 6,530 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses.

Table 3.19: Reduced-Form OLS Regressions of Circumstances on Per Capita Income, National Level

Circumstances	for all HH members				for HHH/spouse only				for HHH/spouse , brothers/sisters, sons/daughters			
	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)
Ethnic Minority / Dominated Caste	**-0.201*** **	**-0.201*** **	**-0.217*** **	**-0.216*** **	-0.204***	-0.204***	-0.213***	-0.213***	-0.195***	-0.195***	-0.211***	-0.211***
	(0.034)	(0.034)	(0.032)	(0.032)	(0.036)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)	(0.033)	(0.033)
Buddhist	0.286***	0.287***	0.302***	0.302***	0.242***	0.242***	0.247***	0.247***	0.253***	0.253***	0.267***	0.267***
	(0.048)	(0.048)	(0.047)	(0.047)	(0.052)	(0.052)	(0.051)	(0.051)	(0.049)	(0.049)	(0.048)	(0.048)
Other Religion	-0.232***	-0.230***	-0.234***	-0.233***	-0.203***	-0.203***	-0.205***	-0.205***	-0.235***	-0.234***	-0.235***	-0.235***
	(0.047)	(0.047)	(0.047)	(0.047)	(0.049)	(0.049)	(0.049)	(0.049)	(0.047)	(0.047)	(0.047)	(0.047)
Father's Education: Primary Complete or Above	0.303***	0.301***	0.304***	**0.302*** **	0.334***	0.334***	0.334***	0.334***	0.328***	0.327***	0.330***	0.329***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.035)	(0.035)	(0.035)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)
Mother's Education: Primary Complete or Above	0.406***	0.413***	0.410***	**0.416*** **	0.372***	0.373***	0.375***	0.375***	0.402***	0.405***	0.404***	0.407***
	(0.045)	(0.045)	(0.045)	(0.045)	(0.053)	(0.053)	(0.054)	(0.053)	(0.047)	(0.047)	(0.047)	(0.046)
Rural Birth Region	**-0.529*** **	**-0.531*** **	**-0.543*** **	**-0.545*** **	-0.424***	-0.424***	-0.431***	-0.431***	-0.490***	-0.492***	-0.502***	-0.504***
	(0.044)	(0.044)	(0.045)	(0.045)	(0.053)	(0.053)	(0.054)	(0.054)	(0.048)	(0.047)	(0.048)	(0.048)
Hill Birth Region	0.189***	0.187***			0.178***	0.178***			0.193***	0.192***		
	(0.052)	(0.052)			(0.054)	(0.054)			(0.054)	(0.054)		
Terai Birth Region	0.101*	0.103*			0.109*	0.109*			0.102*	0.104*		
	(0.055)	(0.055)			(0.056)	(0.056)			(0.056)	(0.056)		
Female	**-0.045**		**-0.039**		-0.005		-0.001		-0.023		-0.016	
	(0.027)		(0.027)		(0.029)		(0.029)		(0.028)		(0.028)	
Constant	11.273***	11.251***	11.431***	11.410***	11.152***	11.149***	11.298***	11.297***	11.225***	11.215***	11.384***	11.376***
	(0.074)	(0.072)	(0.056)	(0.054)	(0.080)	(0.078)	(0.064)	(0.061)	(0.076)	(0.075)	(0.058)	(0.057)
Observations	7065	7065	7065	7065	5920	5920	5920	5920	6583	6583	6583	6583
R-squared	**0.105**	0.105	0.102	0.102	0.081	0.081	0.078	0.078	0.099	0.099	0.096	0.096
R-squared (adjusted)	0.104	0.104	0.101	0.101	0.079	0.079	0.077	0.077	0.098	0.098	0.095	0.095
F stat	92.493	104.005	113.018	131.815	57.869	65.129	70.878	82.694	79.386	89.141	96.906	112.946

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 7,065 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 5,920 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 6,583 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.20: Correlation Matric of Circumstances for All Household Members Using Per Capita Consumption Expenditure

	Female	Ethnic Minority / Dominated Caste	Buddhist	Other Religion	Father's Education Incomplete or Above	Mother's Education Incomplete or Above	Rural Birth Region	Hill Birth Region	Terai Birth Region
Female	**1**								
Ethnic Minority / Dominated Caste	**-0.0329** (0.0057)	**1**							
Buddhist	**-0.0134** (0.2608)	**0.204** (0.0000)	**1**						
Other Religion	**0.0026** (0.8287)	**0.1604** (0.0000)	**-0.0838** (0.0000)	**1**					
Father's Education Incomplete or Above	**0.0018** (0.8831)	**-0.1632** (0.0000)	**0.0017** (0.8893)	**-0.0397** (0.0009)	**1**				
Mother's Education Incomplete or Above	**-0.0811** (0.0000)	**-0.1039** (0.0000)	**-0.0049** (0.6793)	**-0.0564** (0.0000)	**0.3338** (0.0000)	**1**			
Rural Birth Region	**0.0382** (0.0013)	**-0.0829** (0.0000)	**-0.016** (0.1781)	**0.0285** (0.0169)	**-0.1665** (0.0000)	**-0.225** (0.0000)	**1**		
Hill Birth Region	**0.0448** (0.0002)	**-0.2067** (0.0000)	**0.0576** (0.0000)	**-0.0385** (0.0012)	**0.0814** (0.0000)	**0.0167** (0.1601)	**-0.0696** (0.0000)		
Terai Birth Region	**-0.0554** (0.0000)	**0.2457** (0.0000)	**-0.1409** (0.0000)	**0.026** (0.0291)	**-0.0533** (0.0000)	**0.0028** (0.8142)	**0.0185** (0.1204)	**-0.813** (0.0000)	**1**

Note: Display all pairwise correlation coefficients. Significance level for each entry is in parenthesis. Data sample constraints: For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Correlation matrices for the other HH composition samples (HHH and spouse; HHH and spouse, brothers and sisters, sons and daughters) are not displayed here, but available upon request.

Table 3.21: Correlation Matric of Circumstances for All Household Members Using Per Capita Income

	Female	Ethnic Minority / Dominated Caste	Buddhist	Other Religion	Father's Education Incomplete or Above	Mother's Education Incomplete or Above	Rural Birth Region	Hill Birth Region	Terai Birth Region
Female	**1**								
Ethnic Minority / Dominated Caste	**-0.0166** (0.1623)	**1**							
Buddhist	**-0.0029** (0.8098)	**0.2081** (0.0000)	**1**						
Other Religion	**-0.0168** (0.1565)	**0.1607** (0.0000)	**-0.0847** (0.0000)	**1**					
Father's Education Incomplete or Above	**0.0009** (0.9406)	**-0.1779** (0.0000)	**0.003** (0.8029)	**-0.0478** (0.0001)	**1**				
Mother's Education Incomplete or Above	**-0.0647** (0.0000)	**-0.0932** (0.0000)	**-0.0054** (0.6479)	**-0.0362** (0.0023)	**0.3452** (0.0000)	**1**			
Rural Birth Region	**0.0505** (0.0000)	**-0.0873** (0.0000)	**-0.0224** (0.0593)	**0.0335** (0.0048)	**-0.1659** (0.0000)	**-0.2173** (0.0000)	**1**		
Hill Birth Region	**0.0443** (0.0002)	**-0.1935** (0.0000)	**0.0759** (0.0000)	**-0.0294** (0.0131)	**0.0704** (0.0000)	**0.0411** (0.0005)	**-0.0745** (0.0000)	**1**	
Terai Birth Region	**-0.0501** (0.0000)	**0.2531** (0.0000)	**-0.1492** (0.0000)	**0.0187** (0.1146)	**-0.0654** (0.0000)	**-0.0159** (0.1814)	**0.0253** (0.0332)	**-0.8203** (0.0000)	**1**

Note: Display all pairwise correlation coefficients. Significance level for each entry is in parenthesis. Data sample constraints: For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Correlation matrices for the other HH composition samples (HHH and spouse; HHH and spouse, brothers and sisters, sons and daughters) are not displayed here, but available upon request.

Table 3.22: Circumstantial Specific Opportunity Share (Partial IORs), National Level

	All HH members (with data constraints)*	**HHH/spouse** (with data constraints)**	**HHH/spouse, sons/daughters, bros/sis** (with data constraints)***
Per capita consumption expenditure (variable *pccons*)			
Mean Logarithmic Deviation	0.247	0.237	0.242
IOp Index Ratio (parametric)	0.263	0.219	0.250
Circumstances			
Ethnic Minority / Dominated Caste	**0.140**	0.100	0.122
Buddhist	0.068	0.024	0.050
Other Religion	**0.167**	0.137	0.155
Father's Education: Primary Complete or Above	**0.146**	0.106	0.180
Mother's Education: Primary Complete or Above	**0.155**	0.101	0.131
Rural Birth Region	**0.135**	0.090	0.115
Hill Birth Region	**0.133**	0.092	0.116
Terai Birth Region	-	-	-
Female	**0.148**	0.112	0.130
Per capita income (variable *pcincome*)			
Mean Logarithmic Deviation	*0.560*	0.561	0.561
IOp Index Ratio (parametric)	*0.109*	0.096	0.109
Circumstances			
Ethnic Minority / Dominated Caste	0.061	0.057	0.063
Buddhist	0.037	0.033	0.037
Other Religion	0.044	0.042	0.046
Father's Education: Primary Complete or Above	0.070	0.070	0.076
Mother's Education: Primary Complete or Above	0.077	0.062	0.075
Rural Birth Region	0.065	0.059	0.066
Hill Birth Region	0.057	0.053	0.058
Terai Birth Region	0.063	0.058	0.065
Female	-	-	-

Note: only circumstances that were statistically significant at least at the 10 percent level were used to compute the partial IORs.

*For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

**For all 5,899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs.

***For all 6,530 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

3.4.2.2. Most Opportunity Deprived Types and Their Profiles

This section tries to add some pragmatism to the already elaborated descriptive statistics, the IOp scalar indices results and the RF-OLS regression analyses. To do so, the tables in this section ranks the mean economic advantage of types of individuals and displays the number of circumstance variables that are suggested to have a negative impact on economic outcome according to the results of the previous analyses. Table 3.26 and Table 3.27 does this for consumption and income, respectively. Subsequently, Table 3.28 (for consumption) and Table 3.29 (for income) dig further into the characteristics of the most opportunity deprived, that is the ones who accumulate all negatively impacting circumstance variables. They also try to highlight if there is a gender or a Birth Region 2 (Mountain, Hill, Terai) difference. In other words, if they have different impacts on the most opportunity deprived types. This is done because the gender and Birth Region 1 are debated as circumstance variables for reasons already elaborated above. But it is thought that maybe, the most

disadvantaged groups may benefit or suffer additionally depending on which of the sub-categories they accumulate.

For the sake of completeness, this section provides the summary tables for the specific types shares for the most opportunity deprived for both consumption and income. Focus, however, lies on the consumption tables. Also, Table 3.23 displays the national level of mean per capita income and consumption as a reminder. Table 3.24 shows the descriptive statistics for the type distribution running the entire dataset. Both tables use all observations in the NLSS III dataset that have complete information on circumstances and on economic outcome. The HH composition and age constraints are not applied. This is to render the ranking as nationally representative as possible. Almost all observations in the dataset are used (almost 30,000 out of roughly 34,000 observations).

Referring to the consumption panel, one can see that over 68% of the population live below 1.50USD/day and 85% on less than 2.00USD/day. Furthermore, individuals who accumulate 4 or more of the most opportunity depriving characteristic are almost "guaranteed" to be amongst the 43% poorest people in the country. Should one accumulate all five of the most opportunity depriving characteristics, then they are "guaranteed" to live below 1USD/day. Over 4.31% of the population (i.e. 1.14 million people) fall into this category. The estimates for income are less dooming but also alarming.

For the poorest 4.31% of the population, who accumulate all five negative circumstance variables, gender and Birth Region 2 (Mountain, Hill Terai) do not appear to have a different impact on the worst off types (Table 3.28 (for consumption) and Table 3.29 (for income)). There is no gender gap, no ecological zone difference and all most opportunity deprived types appear equally bad off. However, the following level analyses that disaggregate the data by Development Region, urban rural areas and by economic population quarters, paint a very different reality. Refer to the relevant sections for more detail.

Furthermore, Table 3.25 shows the paired t-test results for the mean **statistical differences** between groups of people who accumulate different numbers of negatively impacting circumstances. The table illustrates the results using per capita consumption and for the specific circumstantial share breakdowns at the national level analyses. Similar results were obtained for income and for all other three levels of analyses (available upon request). The null hypothesis states that the mean μ(y) per capita consumption (y) of one group of people accumulating a k number of negatively impacting C circumstances $K_c = \{k_0, \dots, k_c\}$, where $k_c \forall \{0, \dots, 5\}$ is equal to the mean per capita consumption level of another. Formally:

$$H_0\colon\ \mu^k(y) = \mu^l(y)$$
$$H_1\colon\ \mu^k(y) \neq \mu^l(y)$$

As can be seen from the table, all results are statistically significant at the 1 percent level. In other words, for all pairwise combinations of groups that accumulate different numbers of negative circumstances, the null can be rejected. The means of per capita consumption from one group compared to another are statistically different.

Table 3.23: National Level Mean Per Capita Consumption and Income

	NRs	USD*	USD/day
Mean per capita income	97,246.31	1,217.10	3.33
Mean per capita consumption expenditure	45,105.77	564.53	1.55

Note: For the author,s purposed reference variables: *pccons* and *pcincome*.
*av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.24: Sample Partition Description, National Level

	All HH members, Income* (WITHOUT data constraints, 29,722 obs.)	**All HH members, Consumption**** (WITHOUT data constraints, 29,630 obs.)
Maximum number of types	288	288
Number of types observed	221	221
Mean number of observations per type	134.49	134.07
Min. number of observations per type	1	1
Max. number of observations per type	2049	2043
Proportion of types with fewer than 5 observations but more than 0.	16.32 (47)	16.32 (47)

Note: *without data constraints, 29,722 obs.
**without data constraints, 29,630 obs.)

Table 3.25: Statistical Difference Between Means of Types Based on the Number of Accumulated Negative Circumstances, Most Opportunity Deprived (National Level Analyses)

	Number of accumulated negative Circumstances					
Number of accumulated negative Circumstances	0	1	2	3	4	5
0	-	-	-	-	-	-
1	7.92***	-	-	-	-	-
2	13.31***	14.32***	-	-	-	-
3	16.78***	25.26***	18.48***	-	-	-
4	19.04***	32.63***	33.46***	19.38***	-	-
5	19.71***	32.68***	30.27***	17.55***	5.35***	-

Note: T-test's (independent sample, unequal variance) t statistic and statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1. Testing the null that the mean level of *pccons* of 2 of the 6 groups that accumulate negative circumstance variables is equal to 0.

Table 3.26: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure

Number of accumulated negative circumstances	Dominated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth region 2 (M/H/T)	Gender	Types that match	**Population share of sample** (total obs.)°	**Estimated share of overall population***	**Cumulative population share** (%)	**Mean advantage**	**Ratio of overall mean** (%)	**USD per day****
5+Gender	✓	✓	✓	✓	✓		✓Female	6	2.21	585,687	-	28,321.55	62.79	0.97
5	✓	✓	✓	✓	✓			6	4.31	1,140,972	4.31	28,373.44	62.90	0.97
5+Gender	✓	✓	✓	✓	✓		✓Male	6	2.10	555,285	-	28,428.17	63.03	0.97
4	✓	✓	✓		✓			6	0.39	103,725	4.70	29,246.03	64.84	1.00
4	✓	✓	✓	✓				5	0.41	109,090	5.11	31,008.66	68.75	1.06
4	✓		✓	✓	✓			12	31.83	8,432,102	36.94	31,562.17	69.97	1.08
3	✓		✓		✓			12	4.33	1,148,125	41.27	33,097.10	73.38	1.13
4	✓	✓		✓	✓			6	1.23	324,587	42.49	35,728.23	79.21	1.23
4		✓	✓	✓	✓			5	0.06	15,201	**42.55**	36,757.08	81.49	1.26
2		✓			✓			3	0.03	7,153	42.58	37,341.32	82.79	1.28
2			✓		✓			9	2.43	643,808	45.01	37,518.55	83.18	1.29
3			✓	✓	✓			11	11.53	3,055,407	56.54	39,724.40	88.07	1.36
3	✓			✓	✓			12	10.88	2,881,937	67.42	42,046.83	93.22	1.44
3		✓	✓	✓				1	0.00	894	67.42	43,068.45	95.48	1.48
3	✓	✓		✓				5	0.32	84,053	67.74	43,178.33	95.73	1.48
3	✓	✓	✓					5	0.14	37,555	67.88	43,319.88	96.04	1.49
3	✓	✓			✓			6	0.49	128,762	**68.37**	43,374.04	96.16	**1.49**
2		✓		✓				3	0.02	4,471	68.38	48,882.11	108.37	1.68
3		✓		✓	✓			5	0.07	17,884	68.45	51,688.37	114.59	1.77
2	✓				✓			12	5.08	1,345,738	73.53	53,275.12	118.11	1.83
3	✓		✓	✓				12	3.42	906,697	76.95	54,343.23	120.48	1.86
2				✓	✓			8	6.94	1,837,536	83.89	54,344.12	120.48	1.86
2	✓		✓					11	1.21	320,116	**85.10**	58,046.30	128.69	**1.99**
2	✓	✓						6	0.20	51,862	85.29	64,380.04	142.73	2.21
1					✓			6	4.99	1,323,384	90.29	68,293.00	151.41	2.34
1		✓						4	0.05	14,307	90.34	69,439.70	153.95	2.38
2	✓			✓				12	2.60	687,623	92.94	70,394.56	156.07	2.41
3		✓	✓		✓			1	0.00	894	92.94	72,397.82	160.51	2.48
2		✓	✓					1	0.00	894	**92.94**	72,397.82	160.51	**2.48**
1			✓					5	0.52	137,703	93.46	74,177.00	164.45	2.54
1				✓				6	1.08	285,243	94.54	80,711.51	178.94	2.77
2			✓	✓				8	0.69	182,412	95.23	86,399.25	191.55	2.96
1	✓							10	2.82	746,639	98.05	98,080.51	217.45	3.36
0								7	1.95	517,729	100.00	108,356.60	240.23	3.72

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail.
*Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012)
**av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)
°Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure.

Table 3.27: Most Opportunity Deprived Types and Their Profile for Per Capita Income

Number of accumulated negative circumstances	Dominated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth region 2 (M/H/T)	Gender	Types that match	**Population share of sample** (total obs.)°	**Estimated share of overall population***	**Cumulative population share** (%)	**Mean advantage**	**Ratio of overall mean** (%)	**USD per day****
5+Gender	✓	✓	✓	✓	✓		✓Male	6	2.11	558,914	-	53,626.06	55.14	1.84
5	✓	✓	✓	✓	✓			6	4.34	1,149,028	**4.34**	53,658.34	55.18	**1.84**
5+Gender	✓	✓	✓	✓	✓		✓Female	6	2.23	590,114	-	53,688.90	55.21	1.84
4	✓	✓		✓	✓			6	1.22	323,582	5.56	61,662.03	63.41	2.11
2		✓		✓				3	0.02	4,457	5.57	61,761.32	63.51	2.12
3		✓	✓	✓				1	0.00	891	5.58	68,774.60	70.72	2.36
3	✓	✓			✓			6	0.48	128,363	6.06	69,521.20	71.49	2.38
4	✓		✓	✓	✓			12	31.80	8,426,504	37.87	70,447.60	72.44	2.42
3	✓		✓		✓			12	4.32	1,145,463	42.19	72,022.30	74.06	2.47
3	✓	✓	✓					5	0.14	37,439	**42.33**	72,064.13	74.10	**2.47**
4		✓	✓	✓	✓			5	0.06	15,154	42.39	76,143.01	78.30	2.61
4	✓	✓	✓		✓			6	0.39	103,404	42.78	77,302.08	79.49	2.65
3			✓	✓	✓			11	11.57	3,066,452	54.35	84,141.30	86.52	2.89
2			✓		✓			9	2.42	641,816	56.78	84,183.91	86.57	2.89
4	✓	✓	✓	✓				5	0.41	108,752	57.19	88,820.90	91.34	3.05
3	✓	✓		✓				5	0.32	83,793	57.50	90,467.97	93.03	3.10
3	✓			✓	✓			12	10.87	2,880,147	68.37	90,793.14	93.36	3.11
2	✓	✓						6	0.20	51,702	68.57	97,427.74	100.19	3.34
3		✓		✓	✓			5	0.07	17,828	68.64	98,976.97	101.78	3.39
3	✓		✓	✓				12	3.41	903,890	72.05	102,640.90	105.55	3.52
2		✓			✓			3	0.03	7,131	72.07	113,512.70	116.73	3.89
2	✓				✓			12	5.07	1,342,464	77.14	117,216.20	120.54	4.02
2				✓	✓			8	6.95	1,841,654	84.09	118,231.00	121.58	4.05
3		✓	✓		✓			1	0.00	891	84.10	124,113.70	127.63	4.26
2		✓	✓					1	0.00	891	84.10	124,113.70	127.63	4.26
1		✓						4	0.05	14,263	84.15	125,802.30	129.36	4.31
2	✓		✓					11	1.20	319,125	85.36	128,713.00	132.36	4.41
2	✓			✓				12	2.59	685,495	87.94	137,286.80	141.17	4.71
1					✓			6	5.00	1,325,527	92.95	148,683.90	152.89	5.10
1			✓					5	0.52	137,277	93.47	155,832.40	160.25	5.34
1				✓				6	1.07	284,360	94.54	159,343.40	163.86	5.46
2			✓	✓				8	0.70	184,522	95.24	164,101.60	168.75	5.63
1	✓							10	2.81	744,328	98.05	204,848.80	210.65	7.02
0								7	1.95	517,910	100.00	218,811.90	225.01	7.50

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail.
*Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012)
**av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)
°Total of 29,722 obs. with 1) information on all circumstances and 2) information on income.

Table 3.28: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure

Number of accumulated negative circumstances	Dominated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth region 2 (M/H/T)	Gender	Types that match	**Population share of sample** (total obs.)°	**Estimated share of overall population***	**Cumulative population share** (%)	**Mean advantage**	**Ratio of overall mean** (%)	**USD per day****
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Male	1	0.21	54,545	0.21	27,595.35	61.18	0.95
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Female	1	0.84	222,650	1.05	27,609.95	61.21	0.95
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Male	1	0.84	223,545	1.89	28,148.88	62.41	0.97
5+Gender	✓	✓	✓	✓	✓		Female	3	2.21	585,687	-	28,321.55	62.79	0.97
5+Gender	✓	✓	✓	✓	✓		Male	3	2.10	555,285	-	28,428.17	63.03	0.97
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Female	1	1.11	295,079	3.00	28,667.43	63.56	0.98
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Male	1	1.05	277,195	4.05	28,817.29	63.89	0.99
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Female	1	0.26	67,958	4.31	29,151.10	64.63	1.00

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail.
*Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012)
**av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)
°Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure.

Table 3.29: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Income

Number of accumulated negative circumstances	Dominated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth region 2 (M/H/T)	Gender	Types that match	**Population share of sample** (total obs.)°	**Estimated share of overall population***	**Cumulative population share** (%)	**Mean advantage**	**Ratio of overall mean** (%)	**USD per day****
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Female	1	1.13	300,405	1.13	51,373.20	113.89	1.76
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Male	1	0.84	222,853	1.97	51,841.89	114.93	1.78
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Female	1	0.26	67,747	2.23	53,150.38	117.83	1.82
5+Gender	✓	✓	✓	✓	✓		Male	3	2.11	558,914	-	53,626.06	118.89	1.84
5+Gender	✓	✓	✓	✓	✓		Female	3	2.23	590,114	-	53,688.90	119.03	1.84
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Male	1	1.06	281,686	3.29	54,282.05	120.34	1.86
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Female	1	0.84	221,961	4.13	56,987.36	126.34	1.95
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Male	1	0.21	54,376	4.34	57,540.04	127.57	1.97

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail.
*Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012)
**av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)
°Total of 29,722 obs. with 1) information on all circumstances and 2) information on income.

3.4.2.3. Robustness Checks: Economies of Scale, Bootstrapping, Multicollinearity, OLS Regression Validity Checks, etc.

Great care is taken to ensure the most coherent, clean and robust inequality of outcome and IOp analyses. To do so, at every stage of the analytical process, a series of measures was taken to check the data for consistency, the methodology for coherence and the statistical tools and results for robustness. A general overview of all the robustness checks that were done at various stages of the analyses follows. There are seven main "categories" of robustness checks. The same checks were applied at all four levels of analyses: national level, Regional Development level, urban and rural area level, and consumption and income by population quarter level.

Note that the robustness checks summaries here are valid for all level analyses. The subsequent three levels of IOp analyses will only elaborate on particularities in the results of their robustness checks. Elements not mentioned are identical or similar to the robustness check results of the national level analyses.

1) **Dataset**: e.g. data cleaning for outliers, missing values (especially Section 3.3.2)

The NLSS III dataset was cleaned, checked for outliers, missing values, measurement and unit errors etc. Also, each analysed population sample was then checked again for implausible irregularities and extreme values. Section 3.3.2 especially elaborates on how the dataset was cleaned in an effort to make it the most suitable for the IOp analyses. The section explains, for instance, the filling in process of parental education for when data was missing, investigates into the low sex ratio issue of the NLSS III dataset.

2) **Dependant variable aggregations**: e.g. data cleaning, outliers, literature guidance comparison, unit conversion checks (Section 2.4 for consumption and Section 2.5 for income)

Refer to Section 2.4 (for consumption) and Section 2.5 (for income) for full details on the dataset preparations, the dependent variable aggregation process and all the validity checks during the process. For instance, various aggregation guidelines were reviewed and replicated for the dependent variable aggregation. Also, an "optimised" aggregation methodology was proposed having weighed the arguments in the literature that are the most suitable for the NLSS III dataset and the Nepalese context for both economic outcome variables. The literature review section on the measure of economic welfare highlights some of the main arguments and trade-offs of each measure, namely consumption and income. Both economic outcome variables are used as two separate panels in the inequality of outcome and IOp analyses in order to check for the results' consistency. This is, because neither income not consumption are perfect measures for economic welfare. Particularly given the low development context of Nepal, both have different advantages and disadvantages.

3) **Inequality in outcome indices:** e.g. various different indices computations with various dependant variables, bootstrapped standard errors (Sections 2.2 and 2.6)

Sections 2.2 highlights the advantages and disadvantages of the different measures of inequality in outcome. Section 2.6 compares the various inequality in outcome indices that were computed using both income and consumption as proxies for economic welfare. The results were compared, their bootstrapped standard errors computed and the pros and cons of each were reiterated from an empirical standpoint. Checking the total inequality in outcome computations precisely at this

stage is important for the subsequent IOp analysis, which uses the "ideal" Mean Log Deviation measure of inequality in outcome as a reference.

4) **Type distribution and circumstance categorical checks and variations:** e.g. low type distribution checks, circumstance categorical variable combination checks (mainly Sections 3.1.4 and 3.3.2.5)

Section 3.1.4 focuses more on the normative debate on where to draw the line between circumstances and efforts, as well as some other challenges and trade-offs in the measurement of IOp. This is important for the choice of circumstance variables. Section 3.3.2.5 then looks into the definitions and adaptations of the categorical circumstance variables to the NLSS III dataset. The most important trade-off that impacts on the econometric IOp analyses being the fine line between including a maximum number of circumstance variables versus a low type distribution. The lower the level of types observed and the lower the type distribution, the more sensitive the IOp analyses are to extreme values. Thus, great care was taken in choosing the circumstances' categorical variables by closely analysing both the descriptive statistics of the circumstantial variables and then their impact on the econometric analyses.

5) **Regressions:** e.g. various regression tests and various regressions (e.g. Link test, *Ramsey test,* Stepwise regressions, residual checks for normality and homoscedasticity, multicollinearity check, variance inflation factor, outlier and leverage checks, Dfbeta test), (RF-OLS regression sections for each level of analyses: Section 3.4.2.3 for the National level analyses, Section 3.5.2.4 for the Development Region level, Section 3.6.2.4 for the urban rural level, Section 3.7.2.4 for the grouping my population quarter level)

Countless regressions were run, numerous categorical variable combinations were tested and a significant number of regression validity tests were done. Table 3.30 summarises the various regression tests that were run for the finally four chosen RF-OLS regressions for both income and consumption. The table describes the tests, states the results when run using the RF-OLS regressions and gives the author's comments on them. Refer to the table for a quick overview.

The Link test results stated that the results are fine. The same goes for the Ramsey test. Stepwise backward regressions were run and the most significant regressions (4 of which) are given in the analyses. The residuals' normal distributions and homoscedasticity was checked and deemed as acceptable. A multicollinearity check between the independent categorical variables was done. Mainly correlation matrices for all population samples were run and their respective statistical significance level are reported. Some variables correlate, e.g. father's and mother's education levels show a positive correlation. Since both are key explanatory variables, however, they were still included into the regressions. Thus, particular care is taken when the interpreting the results. Furthermore, extreme values are checked for plausibility (at all stages of the analyses) and depending necessary adjustments were done. The data was also checked for the presence of observations that have too much leverage on the predicator variables, and here also changes were done accordingly whenever deemed necessary. Lastly, the DF-BETA test was done to test whether the removal of a coefficient affects the results. The variables that were judged as the most appropriate are reported in the regression tables.

6) **IOp indices:** e.g. parametric and non-parametric, scale equivalence checks, axiom satisfaction checks, income and consumption expenditure analyses comparisons (IOp scalar indices and IOp equivalence scale sections for each level of analyses)

The robustness of the scalar IOp indices was verified using various ways. First, both **parametric and non-parametric approaches** were applied in the computation of the indices. The advantages and disadvantages of each approach are discussed in full detail in Section 3.2.3 that elaborates on the calculations of the scalar indices. The results at this level appear consistent within and between population samples. This is also consistent with the literature and across all levels of the IOp analyses.

Furthermore, the indices were computed for **two separate economic outcome panels**, namely per capita consumption and per capita income. Both of these economic welfare measures bear advantages and disadvantages, which were fully discussed in Section 2.1. The results are consistent with the literature and across all levels of the IOp analyses.

Also, the **axiomatic indices properties** were verified on a theoretical level in order to validate the indices. The basic inequality indices axioms were discussed in Section 2.2 (Literature Review in Inequality in Outcome). Section 3.2.3 on the econometric model applied in this dissertation refers to the basic IOp indices axiomatic properties. Overall, the indices satisfy the discussed axiomatic properties.

Moreover, the **variance of income and consumption and its correlation with the IOp indices** is investigated. First, there is a strong positive correlation between income and consumption expenditure. The correlation coefficient is 0.48 (significance level: 0.0000). Figure 3.17 in the dependent variable definition section (3.3.2.4) shows the scatter plot of pccons against *pcincome* with line of best fit (after having dropped values for *pcincome* below 0 and above 2,000,000NRs. for illustrative purpose).

Second, Table 3.32 shows the correlation coefficients between the variance (or standard deviation) of outcome variables (*pccons*, *pcincome*, and both of them combined) against the non-parametric IOL and IOR estimates as well as MLD. The variance of outcome appears to correlate positively with the MLD estimates, which is as expected (Figure 3.18 illustrates the correlation). The variance of outcome appears to correlate positively with the IOL estimates (Figure 3.19 illustrates the correlation). This is valid for when estimates for both the consumption and income panel are combined as well as for the consumption panel individually. However, it is not valid for when looking at the income panel only. No correlation is apparent.

Last, Ibarra and Martinez Cruz (2015) find in their paper that the IOp estimates have a strong positive correlation with the amount of variation in the outcome variable explained by the combination of circumstances (measured by R-squared). They argue that this calls for the need of high quality and large datasets in order to be able to carry out a reliable and solid IOp analysis. Figure 3.20 is the graphical representation of this result.

Additionally, the IOp scalar indices were check for robustness using the **equivalence scale** methodologies proposed by Buhmann et al. (1988), Coulter et al. (1992) and FG (2011). FG (2011) summarise the equivalence scale approach by Buhmann et al. (1988) as a parametric class of scales given the following transformation: $y_{eq} = y/n^{\alpha}$, where y represents total HH economic outcome (i.e. income or consumption), n HH size and α is the equivalence scale parameter with $\alpha \in [0,1]$. In order to test the IOp indices' sensitivity to scale equivalence, two different scale economy parameters are employed. One where α=0.5 and the other where α=0.75. According to FG (2011, p. 644) "different values for α are used to test the sensitivity of the inequality measures to the different assumptions about

equivalence scale". Table 3.31 presents the newly computed IOp and total inequality indices for α=0.5 and α=0.75 for the national level analyses.

Coulter et al. (1992, p. 1) elaborate at great length on the utility of equivalence scale measures to allow for comparability in income (and consumption) distribution where generally "differences in needs are either ignored or it is assumed". Their value-added to the Buhmann et al.'s (1988) benchmark paper on the utility of equivalence scale measures for poverty and inequality indices is that allows to identify the distributional impact of altering equivalence scale parameters. In fact, it is relatively common in much of the inequality and poverty literature, as well as in the IOp literature, to employ some sort of equivalence scale measure to check for result robustness.

Referring to the **population sample that includes all HH members**, the total inequality indices are the same for both equivalence scale parameters and this is consistent across all population samples and for both economic outcome panels.

For the consumption panel and the parametric estimates, the computed values for both parameters are always the same suggesting that the parametric IOL and IOR estimates are robust and valid. As expected, for both parameters the IOp indices estimates increase. This is because, in general, as α decreases, the intra-household scale economies increase. In other words, richer HH have a disproportionately higher increase in scale economies compared to poorer households. For example, using the estimated indices' values, the normal indicators suggest that between 24% to 29% of overall consumption inequality is due to opportunity deprivation. However, when the equivalence scale measures are applied, the share of opportunity deprivation increases from 36% to 38%. This increase is as expected and consistent.

For the parametric approach, there are some minor differences between the equivalence scale parameters, but they are insignificant and small indicating a significant level of robustness. Furthermore, the estimates between α=0.5 and α=0.75 barely change across all population samples.

For the income panel also, the impact of both parameters within one population sample is not much different.

As expected, the equivalence scale measures for the HH head and spouse only restricted sample are lower than for the sample that includes all HH members. This is because the variance of consumption, for instance, in the sample is amplified by the parameters and the variance is greater in the sample that includes all HH members. This again argues for the robustness of the IOp methodology in this study.

When looking at the **HH head and spouse only restricted dataset** and comparing the impacts of the equivalence scale measures to the literature, the findings of this study are consistent. This argument is valid for both economic outcome panels. For instance, total inequality decreases with both equivalence scale parameters. This is consistent with the literature and theory. FG (2011) argue that as the level of intra-HH scale of economies increases, the total level of inequality decreases. Other estimates follow similar trends as those by FG (2011).

7) **Results**: e.g. comparison between different level analyses of this study, and to other empirical results in the literature (corresponding results interpretations sections for all level analyses, at both the descriptive and econometric level)

The inequality in outcome results are compared to the results obtained by the CBS (2011), the WDI (2011) and other literature. For more details refer to Section 2.6.

There are four levels of IOp analyses in this study. All of the level's results' (descriptive statistics, scalar IOp indices, RF-OLS regression and specific circumstantial share (partial IORs)) are first compared within and then between each analyses level. They statistical differences is tested for.

They are then further compared to mainly the results obtained by FG (2011) and some other IOp literature.

Table 3.30: Various Regression Robustness Check Tests, IOp Analyses at the National Level

Test	Description	Results	Remarks
Link test	Link test: if hatsq is significant i.e. P>|t|: below 0.05, and t above 1.96. This is to say that link test has rejected the assumption that the model is specified correctly.	Hatsq is significant. The results are fine.	
Ramsey test	The Ramsey Test: if Prob > F: lower than 0.05, then it is significant and there is a specification error. Try adding an independent variable or make changes to the dependent variable.	The test is significant. There is a specification error, possibly missing variables. "Ramsey RESET test using powers of the fitted values of ln(y) Ho: model has no omitted variables F(3, 7033) = 17.41 Prob > F = 0.0000"	Yes there are certainly missing variables, e.g. parental occupation, effort. R2 indicates that the model is only valid for about 21% of cases in the sample.
Stepwise Regression (Backward selection)	To check for the inclusion of irrelevant variables.	Various reduced regressions are performed.	The most compelling regression results are represented
Residual normality check	Normality of residuals is only required for valid hypothesis testing. The normality assumption assures that the p-values for the t-tests and F-test will be valid.	Residuals follow a normal distribution pattern against density. This is valid for both the Kernel density estimates and the normal density.	
Residual homoscedasticity check	A suitable model should not represent any pattern once the residuals are plotted against the fitted values.	The fitted values are plotted against the residuals and no patterns are apparent.	
Multicollinearity check	Need to check the relationship between the various independent variables.	Despite R-squared being relatively low (0.21) almost all variables are significant at the 1% level. When the number of observations chances (e.g. data constraint in relationship of individuals to the household head) the results remain almost constant.	
Variance inflation factor	If the mean is >10, then the data should be looked at much closer.	The results are fine.	
Extreme values and outliers	There are many ways to check this. The most intuitive is graphically.	The results are fine.	Checked at all levels of the data cleaning process. But again for the regression.
An observation with "too much" leverage on a predicator variable	There are many ways to check this. The most intuitive is graphically. But there are also some mathematical and listing ways to do so.	The results are fine.	Checked at all levels of the data cleaning process. But again for the regression.
DFBETA test	Affect of the removal of a coefficient, measuring its influence	The results are fine. The used variables are relevant.	

Note: Author's computations and comments.

Table 3.31: Equivalence Scale Robustness Checks on the Scalar Inequality of Opportunity Indices, National Level

	Nepal (for all HH members) *	Equivalence Scale Application		Nepal (for HHH/ spouse only)**	Equivalence Scale Application		Nepal (for HHH/ spouse, sons/ daughters, brothers/ sisters)***	Equivalence Scale Application	
		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75
Per capita consumption expenditure (variable *pccons*)									
Total inequality (E0)	0.247	0.272	0.272	0.237	0.222	0.222	0.242	0.261	0.261
Non-parametric estimates									
IOL	0.077	0.097	0.097	0.067	0.057	0.057	0.075	0.086	0.086
IOR	0.320	0.357	0.357	0.285	0.254	0.254	0.311	0.329	0.329
Parametric estimates									
IOL	0.063	0.103	0.100	0.052	0.050	0.044	0.060	0.087	0.084
IOR	0.263	0.378	0.369	0.219	0.227	0.197	0.250	0.335	0.322
Per capita income (variable *pcincome*)									
Total inequality (E0)	0.560	0.587	0.587	0.561	0.514	0.514	0.561	0.567	0.567
Non-parametric estimates									
IOL	0.094	0.108	0.108	0.094	0.062	0.062	0.099	0.097	0.097
IOR	0.169	0.184	0.184	0.167	0.120	0.120	0.176	0.171	0.171
Parametric estimates									
IOL	0.061	0.120	0.125	0.054	0.060	0.057	0.061	0.107	0.110
IOR	0.109	0.205	0.213	0.096	0.117	0.112	0.109	0.188	0.194

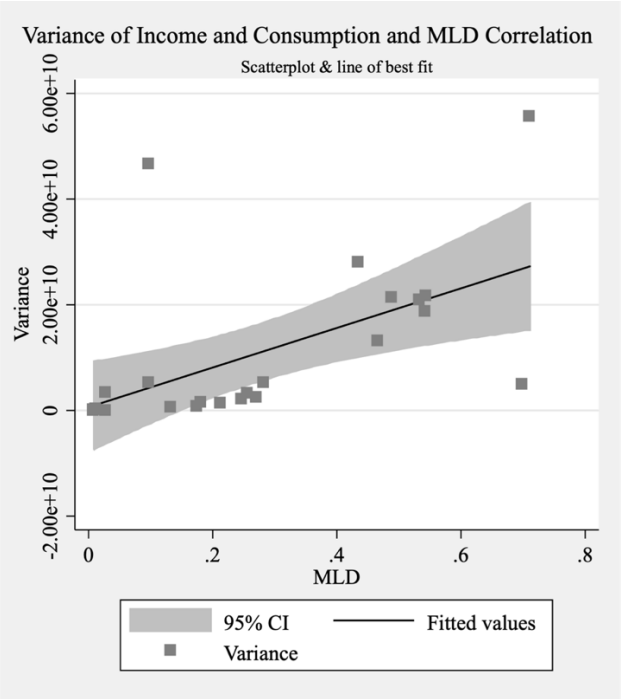
Note: For reference see Buhmann et al. (1998, Coulter et al. (1992).
*For all 7,044 obs. (*pcons*) and 7,065 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have complete information on all circumstance variables.
**For all 5,899 obs. (*pcons*) and 5,920 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have complete information on all circumstance variables.
***For all 6,530 obs. (*pcons*) and 6,583 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have complete information on all circumstance variables.

Table 3.32: Correlation Between the Variance of Outcome Variables and IOR, IOL and MLD

	IOL (non-para)	IOR (non-para)	MLD
Correlation coefficients (combining consumption and income)			
Std. Dev.	0.4673	-0.0048	**0.6893**
(sig. level)	(0.0213)	(0.9823)	(0.0002)
Variance	0.3687	-0.1066	**0.5639**
(sig. level)	(0.0762)	(0.6200)	(0.0041)
Correlation coefficients (consumption only)			
Std. Dev.	**0.7085**	0.6624	**0.7444**
(sig. level)	(0.0099)	(0.0189)	(0.0055)
Variance	0.5931	0.5301	0.5729
(sig. level)	(0.0421)	(0.0763)	(0.0515)
Correlation coefficients (income only)			
Std. Dev.	0.2916	-0.0609	0.5334
(sig. level)	(0.3578)	(0.8509)	(0.0741)
Variance	0.2205	-0.1056	0.3908
(sig. level)	(0.4910)	(0.7439)	(0.2091)

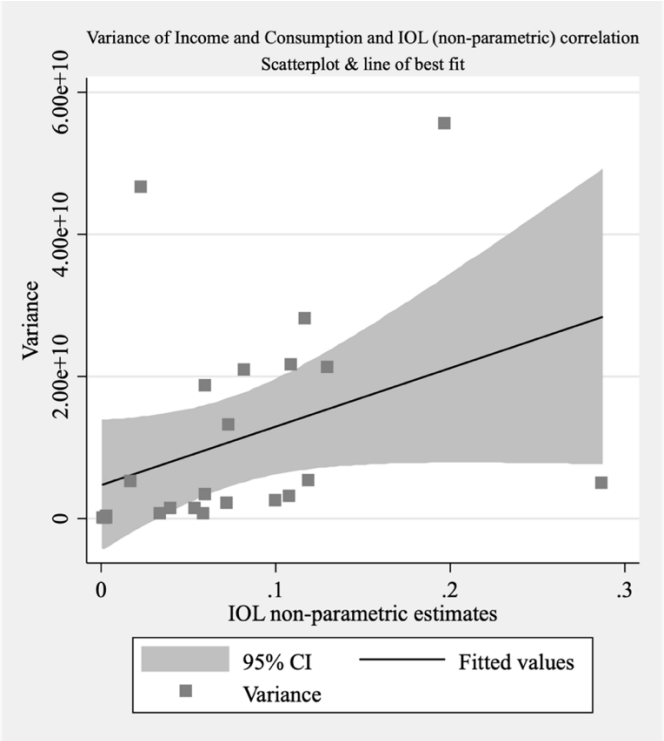
Source: Author's computations. *Note*: Variables *pccons* and *pcincome* serve as reference variables. The estimates are taken from the national level, Development Region level, urban rural area level, and the consumption/income by population quarter level analyses.

Figure 3.18: Variance of Income and Consumption and MLD Correlation



Source: Author's computations. *Note*: Variables *pccons* and *pcincome* serve as reference variables. The estimates are taken from the national level, Development Region level, urban rural area level, and the consumption/income by population quarter level analyses.

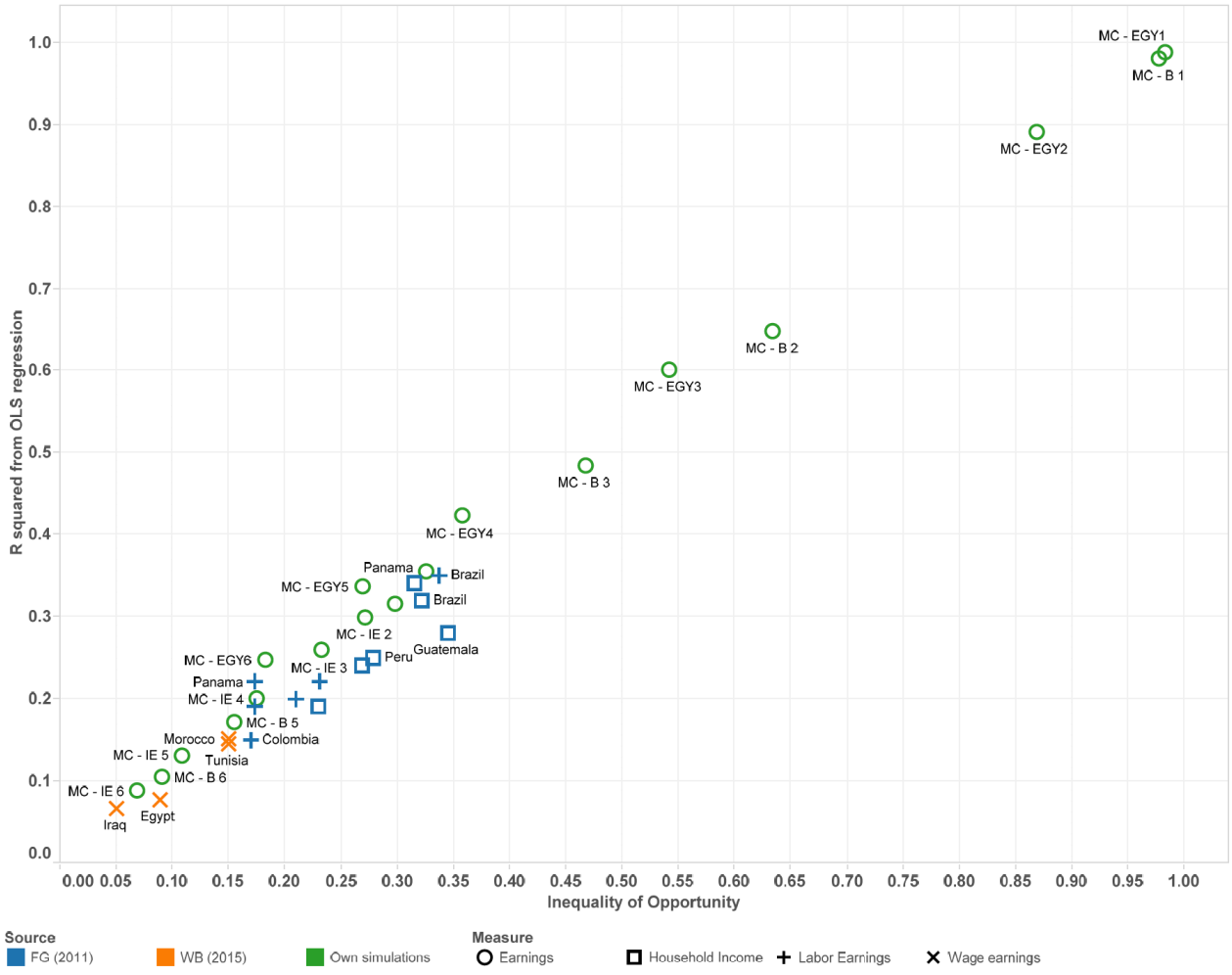
Figure 3.19:Variance of Income and Consumption and IOL Correlation



Source: Author's computations. *Note*: Variables *pccons* and *pcincome* serve as reference variables. The estimates are taken from the national level, Development Region level, urban rural area level, and the consumption/income by population quarter level analyses.

Figure 3.20: Correlation Between Overall Variation Explained and the IOp Estimates (Ibarra and Martinez Cruz, 2015)

Figure 4. Correlation between overall variation explained and the inequality of opportunity estimate



Source: Authors' compilation using data from this study (LIMC, 2015), Ferreira and Gignoux (2011) and World Bank (2015) study in inequality in the MENA region. Notes: The results from LIMC (2015) are obtained from the Monte Carlo simulations in Table 3's Baseline scenarios (labeled MC – B) and Inequality enhanced scenarios (MC – IE) and Table 6's baseline scenario (MC - EGY).

Source: Ibarra and Martinez Cruz (2015, p. 36)

3.5. Inequality of Opportunity Analysis in Nepal: Development Region Level

This section (1.1) focuses on the empirical analyses of IOp at the Development Region level in Nepal. First, this section elaborates on the descriptive statistics.

Second, the empirical results are given and interpreted. To do so, initially the scalar indices of IOp, then the RF-OLS results and the partial IORs are given discussed. Following the circumstantial specific shares of unequal opportunities for the most opportunity deprived are highlighted and discussed. The descriptive and empirical findings of the Development Region level analyses are combined and fully discussed.

Finally, various robustness checks are carried out and discussed.

3.5.1. Descriptive Analysis: Regional Development Level

This section looks at some general noteworthy descriptive statistics that are different to the ones already elaborated and explained in the descriptive statistics section for the national level analyses. It then discusses the type divisions and distributions for the Development Region analyses. Finally, the descriptive statistics for circumstance variables and their sub-categories are given with an initial attempt of interpretation.

The descriptive statistics are given for the five Development Regions: eastern, central, western, mid-western and far-western regions. Note that all HH members who fall into the 30 to 49 age range and have complete information on all circumstances are included into the analyses. This is because the Development Region analyses is Nepal specific and there is limited value-added in constraining the HH composition for the analysis.

The results for per capita income are almost identical to those of per capita consumption and are thus not reported in the tables (available upon request). Only the descriptive type distribution table presents both income and consumption results.

Table 3.35 shows that about 34% of Nepal's population lives in the central region, while the relative population declines towards the further away regions such as the far-west region with only about 9% of Nepal's population. The relative statistics remain constant when the data sample constraints for this study are applied. Since the **sample observations** for the mid-western and far-western regions are relatively low, they are reflected by equally low type distributions (Table 3.37). For instance, the centre observes over 128 out of the possible 288 types, while the far-west only observes about 60 types. Also, more than 62% of these type observations have fewer than 5 observations, while in the centre it is only 42%.

The practice of **joint families** is still common across all Development Regions. This is suggested by the high level of presence of daughters-in-law (see Table 3.36). Also note, that in relative terms, the further away a region is from the central region, which hosts the capital and is the economic and political centre of Nepal, the lower the presence of sons aged between 30 to 49 years. For instance, in the far-west, only 3.8% of HH members are sons and daughters, while in the centre it is over 9.6%. Also, the far-west had the highest proportion of brothers and sisters as well as brothers- and sisters-in-law living within HHs. This suggests both domestic and international mobility. Young adults from the remote rural regions are the most likely to migrate. The centre is the region with the highest rates of HH absorption also of grandchildren, nephews and parents-in-law. This suggests a divergence in living standards and costs of living where a) rural family relatives rely on family networks to try to survive in the urban centres when they migrate, and b) urban centre family members are prone to live with their family up to a higher age until they can afford to live by themselves or with their own families.

This argument is further supported by much smaller than national **sex ratio**, which decreases the further a region is away from the centre. For example, the centre has a sex ratio above the sample

average and equal to the nationally representative average of 89, while the far-west is facing severe male outmigration with a sex ratio of below 65. The bordering eastern region also serves as a hub for immigration with a sex ratio close to 88. This can be a) because it is easy enough to commute to the central region on a regular basis and b) because living expenses there are lower to those of the centre.

It is striking to see from Table 3.33 that only 38% of the people in the far-west (contrary to the national average of over 69%) belong to the dominated caste groups. Again, this suggests that especially the most opportunity deprived are forced to emigrate, because the further away regions from the centre are the most impoverished with the least economic opportunities. Also, these provinces have the smallest shares of religious minorities and urban centres (2% and 5%, respectively, compared to the centre of 5% and 23%).

It is also no surprise that the parental education level of the further away regions is the lowest for the country where over 91% of mothers and 74% of fathers have no schooling, are illiterate or have not completed primary education (versus 60% and 83% of parents in the centre).

Over 18% of people in the central region are **Buddhist**, which is more than double the national average. This is largely due to the proximity of the most densely populated areas of Tibet (e.g. relative proximity to Tibet's capital Lhasa) and to the country of Bhutan. Both of these having been under political turmoil and pressure from China and India with their populations seeking refuge in Nepal, and especially in the eastern region.

Also, over 15% of individuals in the eastern region are religious minorities (double the national average). The main explanation is that the proportion of individuals with indigenous religious beliefs (e.g. Kirats, Jainis, Bons, or Prakritis) in the country historically originate from the eastern region (CBS, 2011; Höfer, 1979).

On a side note, the west has the smallest share of people being born in the mountain belt. This is mainly because the region only has two small provinces that lie in the mountain belt. This is in contrast to the mid-west or far-west where almost 30% to 40% of their territory lie there (refer to Figure 3.9 for a graphical representation and CBD, 2011, p. 6, for the official administrative boundaries).

Table 3.33: Definition and Description of Circumstance Variables, Development Region Level Analysis, (Variable *pccons*)

	Eastern Region, Descriptive stats (%), all HH members*	**Central Region,** Descriptive stats (%), all HH members**	**Western Region,** Descriptive stats (%), all HH members***	**Mid-Western Region,** Descriptive stats (%), all HH members°	**Far-Western Region,** Descriptive stats (%), all HH members°°
Gender					
Dominant (Male)	46.76	47.12	40.08	41.44	39.52
Dominated (Female)	53.24	52.88	59.92	58.56	60.48
Ethnicity					
Dominant caste/ethnic majority	28.24	23.07	38.77	35.03	**61.51**
Dominated caste/ethnic minority	71.76	76.93	61.23	64.97	**38.49**
Religion					
Hindu	77.63	76.72	91.25	96.35	98.11
Buddhist	6.72	**18.19**	5.1	0.22	0.34
Other	**15.65**	5.08	3.65	3.43	1.55
Father's education					
No schooling, illiterate, primary incomplete	64.79	60.31	61.43	**76.02**	**74.4**
Primary completed and above	35.21	39.69	38.57	23.98	25.6
Mother's Education					
No schooling, illiterate, primary incomplete	88.69	83.1	85.26	**90.94**	**91.41**
Primary completed and above	11.31	16.9	14.74	9.06	8.59
Birth region 1 (U/R)					
Urban	8.25	**23.03**	6.61	5.19	5.33
Rural	91.75	**76.97**	93.39	94.81	94.67
Birth region 2 (M/H/T)					
Mountain	10.94	10.29	**0.14**	11.71	20.27
Hill	45.11	60.79	**74.38**	57.13	54.98
Terai	43.95	28.92	25.48	31.16	24.74

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).
Note 2: Parental education was filled using the educational statistics for HHH/spouse brothers/sisters and sons/daughters.
Note 3: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).
*For all 1,636 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 2,479 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
***For all 1,452 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°For all 905 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°°For all 582 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.34: Sex Ratio, Development Region Level Analyses, (Variable *pccons*)

Sex ratio				
Eastern Region, all HH members*	**Central Region,** all HH members**	**Western Region,** all HH members***	**Mid-Western Region,** all HH members°	**Far-Western Region,** all HH members°°
87.83	**89.11**	66.89	70.77	**65.34**

Note: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).

Table 3.35: Birth Place Descriptive Statistics, Development Region Level (Variable *pccons*)

Birth place Development Region	**Descriptive stats (%) all HH members** (no data constraints 29,722 obs.)	**Descriptive stats (%), all HH members** (with data constraints 7,044 obs.)*	**Descriptive stats (%), HHH/spouse** (with data constraints 5,899 obs.)**	**Descriptive stats (%) HHH/spouse, sons/daughters, bros/sis** (with data constraints 6,530 obs.)***
Eastern	23.05	23.34	23.69	23.59
Central	33.52	**35.30**	34.04	35.12
Western	20.03	20.47	20.5	20.44
Mid-western	14.18	**12.75**	13.32	12.76
Far-western	9.22	**8.14**	8.44	8.09
Total	**100.00**	**100.00**	**100.00**	**100.00**

Note: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).

Table 3.36: Relationship to Household Head, Development Region Level Analyses (Variable *pccons*)

	Eastern Region, all HH members*			**Central Region,** all HH members**			**Western Region,** all HH members***			**Mid-Western Region,** all HH members°			**Far-Western Region,** all HH members°°		
Relationship to HH head	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**
Household head	827	50.55	50.55	1,092	44.05	44.05	719	49.52	49.52	464	51.27	51.27	290	49.83	49.83
Husband/Wife	573	35.02	85.57	893	36.02	80.07	485	33.4	82.92	331	36.57	87.85	207	35.57	85.40
Son/Daughter	114	6.97	92.54	238	9.6	89.67	120	8.26	91.18	42	**4.64**	92.49	22	**3.78**	89.18
Brother/Sister	25	1.53	94.07	49	1.98	91.65	22	1.52	92.70	9	0.99	93.48	13	**2.23**	91.41
Father/Mother	3	0.18	94.25	20	0.81	92.46	4	0.28	92.98	7	0.77	94.25	4	0.69	92.10
Grandchild				4	0.16	92.62									
Nephew/Niece				3	0.12	92.74				2	0.22	94.48			
Son/Daughter-In-Law	78	**4.77**	99.02	143	**5.77**	98.51	85	**5.85**	98.83	35	**3.87**	98.34	31	**5.33**	97.42
Brother/Sister-In-Law	8	0.49	99.51	24	0.97	99.48	9	0.62	99.45	13	1.44	99.78	12	**2.06**	99.48
Father/Mother-In-Law				1	0.04	99.52							1	0.17	99.66
Other Family Relative	7	0.43	99.94	8	0.32	99.84	7	0.48	99.93				2	0.34	100
Servant/Servants Relatives				2	0.08	99.92				1	0.11	99.89			
Other Non-Related	1	0.06	100	2	0.08	100	1	0.07	100	1	0.11	100			
Total	**1,636**	**100**		**2,479**	**100**		**1,452**	**100**		**905**	**100**		**582**	**100**	

Note: The descriptive statistics are for the variable *pccons*. The table for *pcincome* can be obtained from the author on request. Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011). Parental education was filled using the educational statistics for HHH/spouse brothers/sisters and sons/daughters. *For all 1,636 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.**For all 2,479 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.***For all 1,452 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °For all 905 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°For all 582 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.37: Type Division Description, Development Region Level Analyses

	Per capita consumption expenditure (*pccons*)					**Per capita income (*pcincome*)**				
	Eastern Region, all HH members*	**Central Region,** all HH members**	**Western Region,** all HH members***	**Mid-Western Region,** all HH members°	**Far-Western Region,** all HH members°°	**Eastern Region,** all HH members*	**Central Region,** all HH members**	**Western Region,** all HH members***	**Mid-Western Region,** all HH members°	**Far-Western Region,** all HH members°°
Maximum number of types	288	288	288	288	288	288	288	288	288	288
Number of types observed	119	128	85	73	**60**	119	128	85	73	60
Mean number of observations per type	13.68067	19.29	17.04	12.32	9.70	13.75	19.37	17.08	12.40	9.70
Min. number of observations per type	1	1	1	1	1	1	1	1	1	1
Max. number of observations per type	166	186	197	133	93	166	186	197	134	93
Proportion of types with fewer than 5 observations but more than 0.	20.49 (59)	19.44 (56)	16.32 (47)	15.28 (44)	12.85 (**37**)	20.49 (59)	19.44 (56)	16.32 (47)	15.28 (44)	12.85 (37)

Note: *For all 1,636 obs. (for both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 2,479 obs. (for both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ***For all 1,452 obs. (for both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °For all 905 obs. (for both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°For all 582 obs. (for both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.5.2. Econometric Analysis for Inequality of Opportunity, Results and Results' Interpretation: Regional Development Level

This econometric analysis section consists first of the scalar indices of IOp. Second, it lays out and interprets the results of the RF-OLS regressions of the circumstance variables on economic outcome. It also looks closely at the circumstantial specific shares of unequal opportunities, the partial IORs. Third, the most opportunity deprived types are ranked by their average mean economic outcome and their profiles are interpreted. Last, the section highlights some of the robustness checks that were carried out throughout the analysis.

3.5.2.1. Parametric and Non-parametric Scalar Indices of Inequality of Opportunity

Table 3.38 presents the results for the scalar IOp indices for the Development Region level analyses. As expected, the total inequality in outcome (MLD or E_0) are higher for income than for consumption. When interpreting the IOp results it is important to refer to the descriptive statistics from the previous section. The sample populations for the mid- and far-western regions are quite small (905 and 582 observations, respectively). This impacts on the econometric results in two main ways. First, it leads to relatively higher levels of total inequality in outcome, and all the more so for income than consumption for reasons previously discussed. Second, the relatively small sample size for the mid- and far-west result in very low type distributions, low type observations and a relatively high proportion of types with less than 5 observations (Table 3.36 and Table 3.37), which in turn affects the IOp estimates. It is for these reasons that the mid-west IOR estimates for income are higher than for consumption, contrary to all other regions, where the consumption estimates are higher than the ones for income. This is an exception, and it is advised to mainly refer to the consumption panel for the IOp Development Region analyses.

As mentioned before, generally the higher levels of IOR consumption to income estimates are due to a higher degree of within type residual inequality in income. This suggests that for the mid-west (like for Guatemala in FG, 2011), the within type components for total inequality in consumption decomposition are, higher. In other words, the IOR(consumption) is lower, because the residual inequality in the consumption distribution is a lot higher, that is there is higher measurement error and transitory consumption variance.

Furthermore, the equivalence scale robustness checks in Section 3.5.2.4 also show that the mid-west indices are very sensitive to scale measures. Again, this can be due to the a) the relatively small sample size of mid-west, and b) a relatively large variance in income distribution in the region.

The highest level of inequality in consumption is attained in the central and western regions. The central region faces the highest level of opportunity deprivation reaching from 37% to 44% of total inequality in consumption expenditure. Unequal opportunities in the mid-west, for instance, "only" make up between 17% to 29%, which is much lower to the national level analyses (26% - 32%).

For the income panel, it is advised to only refer to the IOp indices of the eastern, central and western region and ignore the others due to the reasons elaborated above. These results also suggest that particularly the central region faces the highest share of opportunity deprivation out of overall income inequality (19% - 28%). These estimates are also much above the national level analyses results of 10% to 17%.

The **statistical differences** between the **IOL versus IOR, parametric versus non-parametric estimates** were tested. Table 3.91 shows all the results using the paired t-test. In all cases the null can be rejected and the means of the IOp estimates, no matter whether IOL, IOR, parametric or non-parametric, are different from each other.

Furthermore, the statistical differences between the IOL and IOR estimates (parametric and non-parametric, for income and consumption) **between all levels of the IOp analyses** were tested**,** i.e. national, Development Region, urban rural area, and by income/consumption population quarter level. Table 3.95 reports the results of the paired t-test to test for **statistical difference** between all levels of the IOp analyses. The correlation matrix is available upon request. The coloured boxes within the table show the within level results statistical difference comparison. At the within Development Region level, the computed IOp indices are statistically different between the eastern, central and western regions. There is no statistical difference between the IOp estimates with the far- and mid-western regions.

Overall, the Development Region IOp analyses suggest that people in the central region face the highest level of opportunity deprivation. Combined with the descriptive statistics, which suggest for domestic migration towards the urban areas and the central region, one starts to wonder as to why people migrate to the centre if that is where opportunities are amongst the most unequal. Further interpretations are given in the later sections and particularly in Section 3.8, which compares and interprets the key results across all four different levels of analyses.

Table 3.38: Scalar Indices of Inequality of Opportunity, Development Region Level Analyses

	Eastern Region, all HH members*	**Central Region,** all HH members**	**Western Region,** all HH members***	**Mid-Western Region,** all HH members°	**Far-Western Region,** all HH members°°
Per capita consumption expenditure (variable *pccons*)					
Total inequality (E_0)	0.179	**0.254**	0.267	0.137	**0.184**
Non-parametric estimates					
IOL	0.064	0.112	0.105	0.039	0.071
IOR	0.357	**0.441**	0.393	0.285	0.388
Parametric estimates					
IOL	0.030	0.093	0.085	0.024	0.059
IOR	0.183	**0.365**	0.319	0.171	0.269
Per capita income (variable *pcincome*)					
Total inequality (E_0)	0.484	**0.490**	0.573	0.632	**0.837**
Non-parametric estimates					
IOL	0.103	0.137	0.100	0.201	0.197
IOR	0.213	**0.280**	0.175	**0.318**	0.236
Parametric estimates					
IOL	0.050	0.092	0.067	0.072	0.077
IOR	0.103	**0.188**	0.116	**0.114**	0.091

Note: *For all 1,636 obs. (*pcons*) and 1,636 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 2,469 obs. (*pcons*) and 2,464 obs. (*pcincome*) obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
***For all 1,448 obs. (*pcons*) and 1,423 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°For all 899 obs. (*pcons*) and 894 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°°For all 582 obs. (*pcons*) and 580 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.5.2.2. Reduced-Form OLS Regressions and Circumstance Specific Shares of Opportunities (Partial IORs)

The **RF-OLS regression** results for the Development Region level analyses on per capita consumption expenditure are reported in Table 3.39 and on per capita income in Table 3.40. For comparison purpose to the national level analyses, the same regression and categorical variable combinations are run and reported here. Others are available upon request from the author.

The correlation matrices are not displayed here, but are not greatly different from what is presented in the national level analyses (Table 3.20 and Table 3.21). They are available upon request.

Table 3.41 shows the ***partial IOR*** estimates, that is the circumstance specific shares of overall opportunities compared to overall inequality in outcome. It is important to note that the western region estimates for the consumption panel are prone to error and should be disregarded, instead the income panel estimates should be referred to (for the west only). Overall, the partial IOR results suggest that the circumstances used in this analysis are most adequate for the central region. There the explanatory degree of the circumstances for the IOR is the highest compared to the other regions. In other words, the partial IORs explain most of the opportunity and overall inequality in outcome differentials. It is also the central region that faces the highest share of opportunity deprivation out of overall outcome inequality (36% to 44%), which in turn impacts on the explanatory amplitude of the circumstances.

Overall, no RF-OLS coefficient changes drastically from one regression to another within an analysed sample. R-squared also does not change significantly. Most coefficients have the expected signs and most are statistically significant at the 1% significance level. Overall, the results of the consumption and of the income panel are relatively similar in terms of statistical significance, however, the R-squared value for the consumption panel is much higher across all regressions and regional samples. Overall, the consumption regression tables are preferred for the result interpretations.

On the whole, when comparing the regression results of all five Development Regions, one can conclude that the impact of circumstances on economic outcome diverges greatly among the regions. This goes in contrast with the relatively moderate picture that is painted by the national level analyses. It also highlights the importance of the complementarity and value-added of the Development Region to the national level analyses.

For instance, negatively impacting circumstances on consumption have a much higher amplitude in the central and western region than in the other regions (the statistical difference between them is significant). The level of **ethnic and caste group** discrimination, for example, in central region is about 28% higher than in the far-west. This goes in hand with the national level regression analyses that indicated higher levels of importance for urban and hill area discrimination. The negative impact of dominated caste group belonging on consumption in all other regions, apart from the mid-west, is also quite high. One possible reason for the relatively lower amplitude and significance for the mid-west may be due to the fact that the region is the most homogenous in terms of economic welfare. Everyone is more or less equally bad off and poor (refer to Table 3.80 in the IOp analyses section by economic welfare population grouping). Furthermore, the ethnic minority partial IOR indicates that out of overall consumption inequality up to 23% can be explained by ethnic minority or discriminated caste group belonging. The importance of caste discrimination is by far the highest in the central region.

The coefficient for being **Buddhist** is only statistically significant for the eastern, central and western regions. This is because the proportion of Buddhists living in the mid- and far-west is extremely low (less than 1%, referring to the descriptive statistics in Table 3.33). The highest share of Buddhists live in the central region with over 18%, yet the coefficient is only statistically significant

for the first two regression. Its amplitude is low and negative. A possible explanation is that a) living costs in the centre are much higher than elsewhere, b) the donations received have to suffice for a larger population of monks, and c) the monasteries in the centre have less land and especially agricultural land in order to carry out subsistence farming, so they must purchase more food from the market. These explanations are for the consumption panel estimates. They go in line with the income panel estimates, where the central region is the only region for which the Buddhist coefficient is statistically significant and positive. The monasteries in the centre are forced to diversify their income sources. So some, for instance, provide seasonal Buddhist teaching and mediation classes for tourists (interview with Thubten Jikdol, Lead Monk at Kopan Monetary, Kathmandu, 7 October 2018). The Kopan Monastery is one of the most famous of such monasteries.

Belonging to a religious minority group other than Buddhism has a particularly negative impact on consumption in the central region. This goes in hand with the results from the national level analyses. The coefficients for the eastern region are not significant, possibly due to the religious (indigenous) minorities making up over 15% of believers there and having their own century long enrooted networks there.

The **parental education** coefficients for the Development Region analyses are very interesting. First, for all regions apart for the mid-west, father's education seems more important for enhancing opportunities than mother's education. The statistical difference tests between them show that father's education levels are the most important in the central region. This is interesting, because it is the most developed region in the country, arguably with the most liberal and open minded people. However, there are a few possible interpretations. A) The share of people to emigrate abroad tend to be relatively more liberal and educated than the poorest. This means that the most conservative families remain, where fathers are still the main decision makers. This is backed by the fact that the sex ratio is the highest in the centre out of all regions (89 compared to 65 in the far-west, for instance). B) Due to the higher costs of living in the centre, families are economically forced to live together. So, the proportion of nuclear or single mothered HH is relatively small. Therefore, where mothers benefit from emancipation in the more rural regions such as the mid-west due to fathers emigrating, in the centre another male family member will take the father's role (e.g. brother, uncle or grandfather). C) The centre is not just the economic, but also the political hub of Nepal. Nepal is still at an infant stage of the development of its democratic institutions after decades of political and monarchical turmoil. Only in 2015 a federalisation process was launched. With political parties forming and trying to get voters, their discourse becomes more conservative and relies on the traditional aspects of society in order to populise their discourse. So, the more conservative paroles are much more prominent in the centre than the remote rural areas of Nepal where there is a large disconnect to politics. D) The division of the Development Regions are done for administrative purpose and each Development Region includes both rural and urban areas. Furthermore, 85% of Nepal's population is rural. So when running regressions, the rural observations dominate over the urban ones, possibly biasing the results. For these reasons, the results of this section serve as complementary material for those of the urban rural level analyses section (3.6). So, the parental education interpretation results here needs to be looked at cautiously.

Second, the only region where mother's education is more important to fathers, is the mid-west. The mid-west is not only characterised by heavy male outmigration (extremely low sex ratio), but it is also the region where the level of economic welfare is the most homogenous (i.e. everyone is similarly bad off, with similarly bad access to hard and soft infrastructure). It is therefore no surprise that it is in this region, that the level of mother's education prevails over that of father's.

The RF-OLS results for parental education are partially confirmed by the circumstance specific partial IOR results. However, they should not be given too much importance. First, the partial IORs indicate that the share of parental education are important circumstances in explaining opportunities

and they explain combined from 8% (east) to up to 48% (central region) of overall consumption inequality.

Second, father's education appears to be more important than mother's in the mid- and far-west. Contrary to the regression results, however, mother's education is more important in the centre, where mother's education levels explain more than 26% of opportunities (22% for father's).

Overall, when combining the RF-OLS and the partial IOR results one can conclude that a) there is a vast divergence in the prevalence of parental education across the Development Regions, and b) the results are indecisive on whether mother's or father's education levels are more important for economic outcome and as a share of IOR. Nevertheless, it is suggested that father's education still appears to be the most important in the lesser developed regions.

The amplitude and statistical significance of the **rural birth region** variable vary greatly across regions. It is most significant, negative and takes the largest value in the central and western regions. In these regions and particularly in the urban areas of these regions, living costs are relatively high. When people are born in their urban areas, they can greatly benefit from family networks and family housing. However, being born in a rural area in these regions means a relatively low level of economic welfare compared to the urban centres, while these urban centres drain employment activities away from the rural areas. This makes living in the rural areas, and commuting to or potentially living in the urban areas there very costly.

Referring to the **correlation matrices** (available upon request), the negative correlation between rural birth region and paternal education is largest in the central region. In the further away regions mother's education suffers more from being in the rural area than father's. However this is the opposite in the mid-west. This backs the regression results and interpretations for paternal education, where mother's education is more important than father's in the mid-west.

Table 3.39: Reduced-Form OLS Regressions of Circumstances on Per Capita Consumption Expenditure, Development Region Level

Circum-stances	Eastern Region				Central Region				Western Region				Mid-Western Region				Far-Western Region			
	ln(pccons)				ln(pccons)				ln(pccons)				ln(pccons)				ln(pccons)			
Ethnic Minority / Dominat-ed Caste	-0.283*** (0.034)	-0.284*** (0.034)	-0.314*** (0.032)	-0.315*** (0.032)	**-0.352*** (0.035)**	-0.352*** (0.035)	-0.418*** (0.033)	-0.420*** (0.033)	-0.362*** (0.039)	-0.363*** (0.039)	-0.443*** (0.037)	**-0.446*** (0.036)**	-0.084** (0.036)	-0.084** (0.036)	-0.057 (0.036)	-0.056 (0.035)	-0.275*** (0.050)	-0.276*** (0.050)	-0.209*** (0.047)	-0.208*** (0.048)
Buddhist	0.129** (0.061)	0.130** (0.061)	0.157*** (0.060)	0.158*** (0.059)	**-0.110*** (0.038)**	-0.111*** (0.038)	-0.039 (0.035)	-0.039 (0.035)	0.486*** (0.093)	0.487*** (0.093)	0.584*** (0.091)	**0.589*** (0.091)**	0.011 (0.477)	0.035 (0.475)	0.048 (0.454)	0.071 (0.453)	0.029 (0.271)	0.041 (0.271)	0.048 (0.316)	0.061 (0.316)
Other Religion	-0.021 (0.053)	-0.021 (0.053)	0.003 (0.046)	0.003 (0.046)	-0.387*** (0.060)	-0.385*** (0.060)	-0.388*** (0.058)	-0.387*** (0.058)	-0.143* (0.074)	-0.141* (0.074)	-0.229*** (0.076)	**-0.228*** (0.076)**	0.081 (0.088)	0.078 (0.090)	0.089 (0.087)	0.086 (0.089)	-0.267** (0.123)	-0.271** (0.124)	-0.299** (0.120)	-0.305** (0.121)
Father's Education: Primary Complete or Above	**0.236*** (0.036)**	0.235*** (0.036)	0.232*** (0.037)	0.231*** (0.036)	**0.361*** (0.030)**	0.361*** (0.030)	0.365*** (0.030)	0.366*** (0.030)	**0.240*** (0.037)**	0.242*** (0.037)	0.254*** (0.038)	**0.257*** (0.038)**	**0.276*** (0.040)**	0.280*** (0.040)	0.292*** (0.041)	0.297*** (0.041)	**0.370*** (0.061)**	0.370*** (0.061)	0.377*** (0.061)	0.378*** (0.061)
Mother's Education: Primary Complete or Above	**0.199*** (0.055)**	0.202*** (0.054)	0.196*** (0.055)	0.199*** (0.054)	**0.175*** (0.039)**	0.172*** (0.039)	0.171*** (0.040)	0.167*** (0.040)	**0.207*** (0.054)**	0.201*** (0.053)	0.192*** (0.055)	0.181*** (0.054)	**0.344*** (0.059)**	0.330*** (0.058)	0.337*** (0.059)	0.323*** (0.058)	**0.206** (0.101)**	0.199** (0.099)	0.229** (0.098)	0.221** (0.096)
Rural Birth Region	-0.133** (0.059)	-0.134** (0.059)	-0.108* (0.056)	-0.109* (0.056)	**-0.581*** (0.037)**	-0.582*** (0.037)	-0.630*** (0.035)	-0.630*** (0.035)	**-0.752*** (0.093)**	-0.754*** (0.092)	-0.734*** (0.094)	-0.738*** (0.093)	-0.085* (0.047)	-0.091* (0.048)	-0.115*** (0.044)	-0.121*** (0.045)	-0.324*** (0.101)	-0.323*** (0.102)	-0.388*** (0.102)	-0.388*** (0.104)
Hill Birth Region	0.046 (0.048)	0.046 (0.048)			-0.022 (0.044)	-0.024 (0.044)			**-1.089*** (0.110)**	-1.079*** (0.106)			0.103** (0.052)	0.104** (0.053)			-0.047 (0.061)	-0.044 (0.061)		
Terai Birth Region	-0.028 (0.055)	-0.028 (0.055)			-0.165*** (0.045)	-0.167*** (0.045)			**-1.314*** (0.115)**	-1.308*** (0.111)			0.165*** (0.057)	0.164*** (0.057)			0.135* (0.071)	0.139** (0.070)		
Female	-0.015 (0.028)		-0.013 (0.029)		0.022 (0.025)		0.027 (0.025)		0.027 (0.034)		0.045 (0.034)		0.056* (0.032)		0.055* (0.033)		0.031 (0.050)		0.034 (0.050)	
Constant	10.560*** (0.082)	10.553*** (0.082)	10.558*** (0.063)	10.553*** (0.063)	11.301*** (0.064)	11.316*** (0.062)	11.298*** (0.051)	11.314*** (0.048)	12.387*** (0.127)	12.397*** (0.122)	11.254*** (0.104)	11.286*** (0.097)	9.932*** (0.070)	9.970*** (0.069)	10.043*** (0.056)	10.079*** (0.054)	10.268*** (0.116)	10.283*** (0.116)	10.305*** (0.110)	10.325*** (0.106)
Obs.	1628	1628	1628	1628	2469	2469	2469	2469	1448	1448	1448	1448	899	899	899	899	582	582	582	582
R-squared	0.158	0.158	0.155	0.155	**0.372**	0.372	0.364	0.363	**0.301**	0.3	0.282	0.281	0.133	0.13	0.122	0.119	0.202	0.201	0.188	0.187
R-squared (adjusted)	0.154	0.154	0.152	0.152	0.37	0.37	0.362	0.362	0.296	0.296	0.278	0.278	0.124	0.122	0.115	0.113	0.19	0.19	0.178	0.179
F stat	34.153	38.397	41.453	48.336	183.343	206.29	204.965	239.337	141.063	186.25	69.082	78.784	13.553	14.878	16.118	18.453	15.218	17.035	15.911	18.439

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,628 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 2,469 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,448 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption expenditure aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 580 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.40: Reduced-Form OLS Regressions of Circumstances on Per Capita Income, Development Region Level

Circum-stances	Eastern Region				Central Region				Western Region				Mid-Western Region				Far-Western Region			
	ln(pcincome)				ln(pcincome)				ln(pcincome)				ln(pcincome)				ln(pcincome)			
Ethnic Minority / Dominat-ed Caste	-0.245*** (0.069)	-0.249*** (0.069)	-0.234*** (0.066)	-0.236*** (0.067)	**-0.403*** (0.058)**	-0.402*** (0.058)	-0.480*** (0.055)	-0.477*** (0.054)	-0.208*** (0.074)	-0.208*** (0.074)	-0.264*** (0.069)	-0.265*** (0.069)	-0.048 (0.088)	-0.047 (0.088)	-0.054 (0.086)	-0.054 (0.086)	**-0.379*** (0.111)**	-0.375*** (0.111)	-0.321*** (0.098)	-0.321*** (0.099)
Buddhist	0.212* (0.115)	0.219* (0.115)	0.205* (0.112)	0.211* (0.112)	0.042 (0.063)	0.042 (0.063)	0.122** (0.058)	0.121** (0.058)	0.191 (0.161)	0.192 (0.161)	0.260* (0.157)	0.262* (0.157)	0.543 (0.562)	0.556 (0.559)	0.556 (0.543)	0.567 (0.541)	-0.86 (0.736)	-0.905 (0.736)	-0.867 (0.787)	-0.918 (0.785)
Other Religion	-0.241*** (0.067)	-0.240*** (0.067)	-0.248*** (0.069)	-0.249*** (0.069)	**-0.397*** (0.095)**	-0.399*** (0.095)	-0.389*** (0.093)	-0.390*** (0.093)	-0.072 (0.177)	-0.072 (0.177)	-0.132 (0.175)	-0.131 (0.175)	-0.202 (0.152)	-0.203 (0.152)	-0.224 (0.149)	-0.225 (0.149)	**-1.395*** (0.178)**	-1.382*** (0.172)	-1.366*** (0.147)	-1.344*** (0.139)
Father's Education: Primary Complete or Above	**0.137** (0.061)**	0.131** (0.061)	0.139** (0.060)	0.133** (0.060)	**0.356*** (0.054)**	0.356*** (0.054)	0.367*** (0.055)	0.367*** (0.055)	**0.184*** (0.070)**	0.185*** (0.070)	0.194*** (0.070)	0.195*** (0.071)	**0.267*** (0.101)**	0.270*** (0.101)	0.258** (0.104)	0.261** (0.104)	**0.469*** (0.114)**	0.468*** (0.114)	0.474*** (0.113)	0.473*** (0.112)
Mother's Education: Primary Complete or Above	**0.309*** (0.091)**	0.327*** (0.090)	0.310*** (0.090)	0.328*** (0.089)	**0.186*** (0.064)**	0.191*** (0.063)	0.189*** (0.065)	0.193*** (0.064)	**0.376*** (0.094)**	0.375*** (0.095)	0.366*** (0.095)	0.361*** (0.095)	**0.848*** (0.241)**	0.841*** (0.244)	0.874*** (0.248)	0.867*** (0.252)	0.348** (0.159)	0.377** (0.153)	0.356** (0.154)	0.383*** (0.148)
Rural Birth Region	-0.223** (0.104)	-0.231** (0.104)	-0.232** (0.101)	-0.241** (0.101)	**-0.624*** (0.060)**	-0.624*** (0.060)	-0.725*** (0.055)	-0.724*** (0.055)	**-0.714*** (0.137)**	-0.715*** (0.136)	-0.702*** (0.137)	-0.704*** (0.136)	-0.335** (0.139)	-0.338** (0.139)	-0.354*** (0.134)	-0.357*** (0.135)	-0.240 (0.207)	-0.246 (0.214)	-0.293 (0.207)	-0.295 (0.214)
Hill Birth Region	-0.039 (0.090)	-0.04 (0.090)			**0.229*** (0.073)**	0.233*** (0.073)			-0.808*** (0.195)	-0.806*** (0.192)			-0.145 (0.143)	-0.144 (0.143)			-0.154 (0.134)	-0.164 (0.135)		
Terai Birth Region	-0.011 (0.094)	-0.008 (0.094)			-0.013 (0.078)	-0.009 (0.078)			-0.965*** (0.205)	-0.963*** (0.204)			-0.075 (0.155)	-0.076 (0.155)			0.055 (0.156)	0.040 (0.157)		
Female	-0.079 (0.056)		-0.08 (0.056)		-0.036 (0.044)		-0.035 (0.044)		0.007 (0.062)		0.02 (0.061)		0.030 (0.074)		0.027 (0.074)		-0.121 (0.100)		-0.124 (0.100)	
Constant	11.326*** (0.143)	11.292*** (0.140)	11.306*** (0.118)	11.273*** (0.115)	11.804*** (0.107)	11.780*** (0.103)	12.038*** (0.083)	12.017*** (0.077)	12.555*** (0.221)	12.558*** (0.218)	11.716*** (0.154)	11.731*** (0.146)	10.852*** (0.199)	10.871*** (0.192)	10.776*** (0.160)	10.794*** (0.152)	10.836*** (0.246)	10.777*** (0.244)	10.791*** (0.220)	10.720*** (0.217)
Obs.	1631	1631	1631	1631	2467	2467	2467	2467	1423	1423	1423	1423	894	894	894	894	580	580	580	580
R-squared	0.061	0.059	0.06	0.059	**0.208**	0.208	0.196	0.196	0.086	0.086	0.082	0.082	0.08	0.08	0.078	0.077	0.131	0.128	0.125	0.122
R-squared (adjusted)	0.055	0.054	0.056	0.055	0.205	0.205	0.194	0.194	0.08	0.081	0.077	0.078	0.071	0.072	0.07	0.071	0.118	0.116	0.114	0.113
F stat	9.73	10.855	12.188	14.079	84.713	94.78	97.519	112.764	22.463	25.64	17.101	19.968	6.302	6.58	8.001	8.68	20.749	24.773	27.688	34.648

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 1,631 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 2,467 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 1,423 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 894 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 580 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.41: Circumstantial Specific Opportunity Share (Partial IORs), Development Region Level

	Eastern Region (all HH members)°	**Central Region** (all HH members)°°	**Western Region** (all HH members)°°°	**Mid-Western Region** (all HH members) °°°°	**Far-Western Region** (all HH members)° °°°°
Per capita consumption expenditure (variable *pccons*)					
Mean Logarithmic Deviation	0.179	0.254	*0.267*	0.137	0.184
IOp Index Ratio (parametric)	0.183	0.365	*0.319*	0.171	0.269
Circumstances					
Ethnic Minority / Dominated Caste	0.040	0.227	*0.701*	0.051	0.121
Buddhist	0.062	0.139	*0.668*	-	-
Other Religion	-	0.130	*0.666*	-	0.101
Father's Education: Primary Complete or Above	0.017	0.219	*0.614*	0.107	0.130
Mother's Education: Primary Complete or Above	0.060	0.264	*0.661*	0.056	0.128
Rural Birth Region	0.047	0.225	*0.469*	0.051	0.120
Hill Birth Region	-	-	*0.059*	0.016	-
Terai Birth Region	-	0.221	*0.657*	0.051	0.120
Female	-	-	-	0.047	-
Per capita income (variable *pcincome*)					
Mean Logarithmic Deviation	0.484	0.490	0.573	0.632	0.837
IOp Index Ratio (parametric)	0.103	0.188	0.116	0.114	0.091
Circumstances					
Ethnic Minority / Dominated Caste	0.018	0.111	0.272	-	0.026
Buddhist	0.034	0.062	0.269	-	-
Other Religion	0.010	0.071	-	-	0.132
Father's Education: Primary Complete or Above	0.048	0.104	0.241	0.096	0.098
Mother's Education: Primary Complete or Above	0.003	0.135	0.267	0.000	0.098
Rural Birth Region	0.002	0.126	0.222	0.008	-
Hill Birth Region	-	0.111	0.029	-	-
Terai Birth Region	-	-	0.268	-	-
Female	-	-	-	-	-

Note: only circumstances that were statistically significant at least at the 10 percent level were used to compute the partial IORs.

°For all 1,636 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°°For all 2,469 obs. for per capita consumption expenditure and 2,464 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°°°For all 1,448 obs. for per capita consumption expenditure and 1,423 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°°°°For all 899 obs. for per capita consumption expenditure and 894 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°°°°°For all 582 obs. for per capita consumption expenditure and 580 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.5.2.3. Most Opportunity Deprived Types and Their Profiles

This section further builds on the national level results that ranks the types of individuals by mean economic advantage. This section does the same but for the 5 Development Region samples. The tables presented here are for consumption only. The results and trends of the income tables do not vary hugely, and are thus available upon request. Table 3.44 through to Table 3.48 ranks the different types of populations by their mean per capita consumption for each of the five Development Regions. Table 3.49 through to Table 3.53 dig further into the characteristics of the most opportunity deprived by further disaggregating them by gender and Birth Region 2 (Mountain, Hill, Terai).

For the sake of completeness, this section provides summary tables for the mean levels of per capita consumption for each Development Region (Table 3.42) and the descriptive statistics for the type distribution (Table 3.43). Both tables use all observations in the NLSS III dataset that have complete information on circumstances and on consumption. The HH composition and age constraints are not applied. This is to render the ranking as nationally representative as possible. Almost all observations in the dataset are used.

In general and for all regions, the more negative circumstances one accumulates, the more likely one is to have a lower level of consumption expenditure, and the more likely they are to suffer from opportunity deprivation (when combining the ranking results with those of the descriptive and econometric results from the previous sections).

There are some **circumstantial accumulation outliers** due to very low type distributions. The far-west, for instance, observes a type where dominated caste belonging is the only negative circumstance (i.e. not many negative circumstances are accumulated), yet, that type's mean consumption is only 54% of the national mean consumption.

The mid-west and the far-west have the highest share of people living on **less than 1USD/day** (80% and 70%, respectively). All other three regions have below 4% of the population with less than 1USD/day.

A few interesting observations can be made when referring to the tables that complement the **most opportunity deprived types** (i.e. individuals who accumulate all five negative circumstantial variables) with gender and Birth Region 2 characteristics. First, in the **west**, those who belong to the most opportunity deprived type and live in the hill region are an exception. Their level of consumption expenditure (no matter the gender) is around the same as the national mean, while the other four most disadvantaged types only consume between 50% to 60% of the national mean. The hill region in the west hosts the city of Pokhara, a cultural heritage city, that is also extremely popular with outdoor tourists and is at relative proximity to the capital (30 minutes flight). Also, the hill region in the west is known to be a hub for foreign aid and domestic development projects. JICA or the German equivalent GIZ have a series of ongoing projects and representative offices there (various interviews with GIZ and JICA officials, October 2018). Moreover, the Pokhara region hosts two of the largest Tibetan refugee camps in Nepal attracting further attention from the international community (various interviews with representatives of the Tibetan community in Nepal, October 2018). Also, the Pokhara area is a training region for the Gurkha soldiers that use this British military training as a means of

making a living for themselves and leave Nepal[10]. For all these reasons, especially people from the most disadvantaged groups in the western region's hill area benefit from various development programs and are able to withhold a higher level of consumption expenditure.

Second, of the most disadvantaged type, men who live in the mountain belt of the **central region** have a consumption expenditure level that is more than double that of all other most disadvantaged types. A possible explanation of this is that the central region's mountain areas are amongst the most famous hiking destinations in the world and are extremely popular with foreign tourists. It is possible that many of these men work in tourism as luggage carriers etc where they get the chance to work (and receive tips, which a) women can't do, and b) usually they would not get in the region without tourism). Actually, the women in the mountain belt are the poorest of all observed type combinations with less than 32% for the central region's mean consumption level. So there is a significant gender gap between men and women in the mountain belt of the central region. Apart from that, when more than three negative circumstances are accumulated, mean consumption drops below 63% of the national mean.

Third, the **far-west** statistics for the most opportunity deprived types show a clear ecological belt impact on consumption expenditure. Yet, there is no gender gap within the ecological bets within the far-western region. People who accumulate all five negative circumstances and are born in the mountain belt are deemed to live with less than 30% of the national consumption expenditure level (and 48% of the mean level of the far-west), while this is more than 12% higher (or 17%) for people born in the Terai belt within the far-west. A possible explanation is that the far-west is the number one region that is the most dependant on agricultural income (direct agricultural and agricultural wage income; refer to Section 2.6 for more details). The low lying fertile lands of the Terai region are suitable for agricultural and livestock produce, which is of imminent importance for the poorest of the poor. The entire far-west and the mid-west are the poorest two regions in Nepal where over 99% of the population lives below 2 USD/day (consumption panel) and more than 70 and 79% of their populations below 1USD/day. So, if on top of that their populations are mainly only dependent on faming income, then it becomes evident that the further north one moves across the ecological belts (first the hill and then then mountain belt), the lower the potential farm revenues are. This is reflected by results in Table 3.53) for the most opportunity deprived types.

Table 3.42: Development Region Level Mean Per Capita Consumption

Mean per capita consumption expenditure	**NRs**	**USD***	**USD/day**
Eastern Region	40,147.17	502.47	1.38
Central Region	58,779.62	735.66	2.02
Western Region	47,527.52	594.84	1.63
Mid-Western Region	27,881.96	348.96	0.96
Far-Western Region	28,855.48	361.14	0.99
Nepal total	**45,105.77**	**564.53**	**1.55**

Note: For the author's purposed reference variables: *pccons.* *all HH members, all obs. in the Development Region that have information on all circumstances and per capita consumption expenditure
**av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

[10] Extensive documentation can be obtained on the British Royal Army's website, the Brigade of Gurkhas: https://www.army.mod.uk/who-we-are/corps-regiments-and-units/brigade-of-gurkhas/.

Table 3.43: Sample Partition Description, Development Region Level

	Eastern Region, all HH members	**Central Region,** all HH members	**Western Region,** all HH members	**Mid-Western Region,** all HH members	**Far-Western Region,** all HH members
Maximum number of types	288	288	288	288	288
Number of types observed	175	188	123	116	94
Mean number of observations per type	38.99	52.89	48.33	36.03	29.15
Min. number of observations per type	1	1	1	1	1
Max. number of observations per type	555	687	614	473	269
Proportion of types with fewer than 5 observations but more than 0.	14.58 (42)	17.01 (49)	14.58 (42)	13.19 (38)	8.68 (25)

Note: all HH members, all 29,630 obs. in the Development Region that have information on all circumstances and per capita consumption expenditure *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region.

Table 3.44: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Development Region Level, Eastern Region

							Eastern Region								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Gender	**Population share of sample** (total obs.)*	**Cumulative Eastern Region population share** (%)	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Eastern Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
2		✓			✓		0.00	-	-	-	-	-	0.00	0.00	-
2		✓	✓				0.00	-	-	-	-	-	0.00	0.00	-
4	✓	✓	✓	✓			0.81	**0.81**	0.19	0.19	49,180	24,668.49	61.45	54.69	**0.85**
4	✓	✓	✓		✓		1.35	2.15	0.31	0.50	82,264	29,894.71	74.46	66.28	1.03
3	✓		✓	✓			2.86	5.01	0.66	1.15	174,365	30,677.79	76.41	68.01	1.05
5+Gender	✓	✓	✓	✓	✓	✓Female	5.33	-	1.23	-	325,481	31,230.42	77.79	69.24	1.07
5	✓	✓	✓	✓	✓		9.81	14.82	2.26	3.41	598,205	31,572.05	78.64	70.00	1.08
3	✓		✓		✓		4.13	18.95	0.95	4.36	252,158	31,699.94	78.96	70.28	1.09
5+Gender	✓	✓	✓	✓	✓	✓Male	4.47	-	1.03	-	272,724	31,979.77	79.66	70.90	1.10
4	✓		✓	✓	✓		29.49	48.44	6.79	11.15	1,799,087	32,707.96	81.47	72.51	1.12
3	✓	✓		✓			0.60	49.04	0.14	11.29	36,661	32,859.92	81.85	72.85	1.13
3	✓	✓	✓				0.29	49.33	0.07	11.36	17,884	37,528.46	93.48	83.20	1.29
4	✓	✓		✓	✓		3.18	**52.51**	0.73	12.09	194,037	37,706.17	93.92	83.60	**1.29**
3	✓			✓	✓		10.35	62.86	2.38	14.48	631,290	41,206.11	102.64	91.35	1.41
2			✓		✓		2.02	64.88	0.47	14.94	123,397	41,399.82	103.12	91.78	1.42
3			✓	✓	✓		10.11	75.00	2.33	17.27	616,983	41,611.72	103.65	92.25	1.43
3	✓	✓			✓		1.55	76.55	0.36	17.63	94,783	42,886.15	106.82	95.08	1.47
3		✓	✓	✓			0.01	76.56	0.00	17.63	894	43,068.45	107.28	95.48	1.48
2		✓		✓			0.01	76.58	0.00	17.63	894	43,068.45	107.28	95.48	1.48
2	✓			✓			1.54	**78.12**	0.35	**17.99**	93,889	43,123.12	107.41	95.60	**1.48**
2	✓				✓		5.01	83.13	1.15	19.14	305,809	44,387.76	110.56	98.41	1.52
1			✓				0.29	83.42	0.07	19.21	17,884	46,831.03	116.65	103.82	1.61
2	✓		✓				1.03	84.45	0.24	19.45	62,592	47,548.17	118.43	105.41	1.63
2				✓	✓		6.65	**91.10**	1.53	20.98	405,957	51,414.55	128.07	113.99	**1.76**
2	✓	✓					0.40	91.50	0.09	21.07	24,143	61,073.67	152.12	135.40	2.09
1					✓		4.63	96.13	1.07	22.14	282,560	61,244.83	152.55	135.78	2.10
2			✓	✓			0.54	96.67	0.12	22.26	33,085	62,100.92	154.68	137.68	2.13
1				✓			0.44	97.11	0.10	22.36	26,825	63,911.93	159.19	141.69	2.19
1	✓						1.89	99.00	0.44	22.80	115,349	71,921.44	179.14	159.45	2.47
3		✓		✓	✓		0.01	99.02	0.00	22.80	894	72,397.82	180.33	160.51	2.48
3		✓	✓		✓		0.01	99.03	0.00	22.80	894	72,397.82	180.33	160.51	2.48
4		✓	✓	✓	✓		0.01	99.05	0.00	22.81	894	75,487.11	188.03	167.36	2.59
0							0.88	99.93	0.20	23.01	53,651	115,393.10	287.43	255.83	3.96
1		✓					0.07	100.00	0.02	23.03	4,471	116,305.70	289.70	257.85	3.99

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.45: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Development Region Level, Central Region

							Central Region								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Gender	**Population share of sample** (total obs.)*	**Cumulative Central Region population share** (%)	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Central Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~3~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~			~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~	~~✓~~	~~✓~~	~~✓~~		~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
5+Gender	✓	✓	✓	✓	✓	✓Female	1.60	-	0.54	-	142,174	23,445.68	39.89	51.98	0.80
5	✓	✓	✓	✓	✓		3.37	**3.37**	1.13	1.13	299,550	24,338.02	41.41	53.96	0.83
5+Gender	✓	✓	✓	✓	✓	✓Male	1.77	-	0.59	-	157,375	25,144.17	42.78	55.74	0.86
4	✓	✓	✓		✓		0.19	**3.56**	0.06	1.19	16,989	25,837.37	43.96	57.28	**0.89**
4	✓	✓		✓	✓		0.53	4.09	0.18	1.37	47,391	34,448.51	58.61	76.37	1.18
4	✓		✓	✓	✓		33.38	37.47	11.20	12.58	2,967,778	36,915.74	62.80	81.84	1.27
4	✓	✓	✓	✓			0.61	38.09	0.21	12.78	54,545	37,044.43	63.02	82.13	1.27
4		✓	✓	✓	✓		0.06	38.15	0.02	12.80	5,365	37,370.22	63.58	82.85	1.28
3	✓	✓			✓		0.17	38.32	0.06	12.86	15,201	38,002.81	64.65	84.25	1.30
3	✓		✓		✓		4.29	42.61	1.44	14.30	381,814	40,934.38	69.64	90.75	1.40
2		✓			✓		0.06	**42.67**	0.02	14.32	5,365	42,830.89	72.87	94.96	**1.47**
3	✓	✓	✓				0.19	42.86	0.06	14.38	16,989	45,562.77	77.51	101.01	1.56
1		✓					0.11	42.97	0.04	14.42	9,836	48,136.96	81.89	106.72	1.65
2		✓		✓			0.04	43.02	0.01	14.43	3,577	50,335.53	85.63	111.59	1.73
3	✓			✓	✓		10.79	53.81	3.62	18.06	959,453	51,308.34	87.29	113.75	1.76
3	✓	✓		✓			0.42	54.23	0.14	18.20	37,555	52,273.83	88.93	115.89	1.79
3		✓		✓	✓		0.08	54.31	0.03	18.22	7,153	52,524.95	89.36	116.45	1.80
2			✓		✓		1.25	**55.56**	0.42	18.64	110,878	57,679.35	98.13	127.88	**1.98**
3			✓	✓	✓		6.59	62.14	2.21	20.85	585,687	59,394.40	101.05	131.68	2.04
3	✓		✓	✓			5.85	68.00	1.96	22.82	520,412	68,526.13	116.58	151.92	2.35
2				✓	✓		4.57	72.56	1.53	24.35	405,957	68,706.93	116.89	152.32	2.36
2	✓		✓				2.05	74.62	0.69	25.04	182,412	68,997.82	117.38	152.97	2.37
2	✓				✓		4.57	79.18	1.53	26.57	405,957	70,769.97	120.40	156.90	2.43
2		✓	✓				0.01	79.19	0.00	26.57	894	72,397.82	123.17	160.51	2.48
2	✓	✓					0.19	79.38	0.06	26.64	16,989	73,397.76	124.87	162.72	2.52
1			✓				0.68	80.07	0.23	26.87	60,804	81,257.72	138.24	180.15	2.79
2	✓			✓			4.90	84.96	1.64	28.51	435,465	84,829.70	144.32	188.07	2.91
1					✓		4.00	88.97	1.34	29.85	355,883	93,158.73	158.49	206.53	3.19
1				✓			1.58	90.55	0.53	30.38	140,386	95,259.69	162.06	211.19	3.27
1	✓						5.60	96.15	1.88	32.26	498,057	108,910.90	185.29	241.46	3.73
2			✓	✓			0.78	96.93	0.26	32.53	69,746	114,653.90	195.06	254.19	3.93
0							3.07	100.00	1.03	33.56	272,724	124,723.70	212.19	276.51	4.28

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.46: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Development Region Level, Western Region

							Western Region								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no edu-cation	Mother's with no edu-cation	Rural Birth Region	Gender	**Population share of sample** (total obs.)*	**Cumulative Western Region population share** (%)	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~3~~		✓	✓	✓			~~0.00~~	-	-	-	-	-	~~0.00~~	~~0.00~~	-
~~2~~		✓		✓			~~0.00~~	-	-	-	-	-	~~0.00~~	~~0.00~~	-
~~1~~		✓					~~0.00~~	-	-	-	-	-	~~0.00~~	~~0.00~~	-
~~3~~		✓	✓		✓		~~0.00~~	-	-	-	-	-	~~0.00~~	~~0.00~~	-
~~2~~		✓	✓				~~0.00~~	-	-	-	-	-	~~0.00~~	~~0.00~~	-
2		✓			✓		0.02	0.02	0.00	0.00	894	20,872.60	43.92	46.27	0.72
4		✓	✓	✓	✓		0.12	0.13	0.02	0.03	6,259	25,072.86	52.75	55.59	0.86
4	✓	✓	✓	✓			0.02	0.15	0.00	0.03	894	25,999.05	54.70	57.64	0.89
4	✓	✓	✓		✓		0.05	0.20	0.01	0.04	2,683	27,239.72	57.31	60.39	0.93
5+Gender	✓	✓	✓	✓	✓	✓Male	1.30	-	0.26	-	68,852	28,323.97	59.59	62.79	0.97
5	✓	✓	✓	✓	✓		2.81	**3.01**	0.56	**0.60**	149,328	28,451.03	59.86	63.08	0.98
5+Gender	✓	✓	✓	✓	✓	✓Female	1.51	-	0.30	-	80,476	28,559.74	60.09	63.32	**0.98**
4	✓		✓	✓	✓		33.93	36.94	6.81	7.41	1,803,558	31,039.70	65.31	68.82	1.06
3	✓		✓		✓		4.61	41.55	0.92	8.34	245,005	31,450.47	66.17	69.73	1.08
4	✓	✓		✓	✓		1.09	42.65	0.22	8.56	58,122	34,915.12	73.46	77.41	1.20
2			✓		✓		2.54	45.19	0.51	9.07	135,021	41,745.43	87.83	92.55	1.43
2	✓	✓					0.08	45.27	0.02	9.08	4,471	41,930.26	88.22	92.96	1.44
3	✓			✓	✓		12.55	**57.82**	2.52	**11.60**	667,057	43,547.18	91.63	96.54	**1.49**
2	✓		✓				0.79	58.61	0.16	11.76	42,026	46,909.15	98.70	104.00	1.61
3			✓	✓	✓		10.53	69.15	2.11	13.87	559,756	50,742.04	106.76	112.50	1.74
3		✓		✓	✓		0.05	69.20	0.01	13.88	2,683	52,190.99	109.81	115.71	1.79
2	✓				✓		6.26	**75.45**	1.26	**15.14**	332,634	56,367.22	118.60	124.97	**1.93**
3	✓		✓	✓			1.58	77.04	0.32	15.45	84,053	59,766.49	125.75	132.50	2.05
3	✓	✓		✓			0.10	77.14	0.02	15.47	5,365	64,099.42	134.87	142.11	2.20
2	✓			✓			1.55	78.68	0.31	15.78	82,264	65,991.92	138.85	146.30	2.26
2				✓	✓		8.50	87.18	1.70	17.49	451,560	66,917.20	140.80	148.36	2.29
3	✓	✓			✓		0.20	87.38	0.04	17.53	10,730	67,917.44	142.90	150.57	2.33
1				✓			0.98	88.36	0.20	17.73	51,862	71,599.11	150.65	158.74	2.46
1					✓		6.58	94.94	1.32	19.04	349,624	72,606.36	152.77	160.97	2.49
3	✓	✓	✓				0.03	94.97	0.01	19.05	1,788	86,015.57	180.98	190.70	2.95
1			✓				0.72	95.69	0.15	19.20	38,450	92,011.56	193.60	203.99	3.16
2			✓	✓			0.64	96.33	0.13	19.33	33,979	94,629.27	199.10	209.79	3.24
1	✓						1.72	98.05	0.34	19.67	91,206	99,856.52	210.10	221.38	3.42
0							1.95	100.00	0.39	20.06	103,725	109,711.80	230.84	243.23	3.76

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.47: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Development Region Level, Mid-Western Region

											Mid-Western Region				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no edu-cation	Mother's with no education	Rural Birth Region	Gender	**Population share of sample** (total obs.)*	**Cumulative Mid-Western Region population share** (%)	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advan-tage**	**Ratio of Mid-Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
3		✓	✓	✓			0.00						0.00	0.00	
2		✓		✓			0.00						0.00	0.00	
1		✓					0.00						0.00	0.00	
3		✓	✓		✓		0.00						0.00	0.00	
2		✓	✓				0.00						0.00	0.00	
2		✓			✓		0.02	0.02				20,872.60	74.86	46.27	0.72
5+Gender	✓	✓	✓	✓	✓	✓Male	1.32	-	0.19	-	49,180	21,227.36	76.13	47.06	0.73
5	✓	✓	✓	✓	✓		2.18	2.20	0.31	0.31	81,370	21,581.64	77.40	47.85	0.74
5+Gender	✓	✓	✓	✓	✓	✓Female	0.86	-	0.12	-	32,190	22,122.88	79.34	49.05	0.76
3	✓		✓		✓		4.83	7.03	0.68	0.99	180,624	23,277.94	83.49	51.61	0.80
4	✓		✓	✓	✓		35.53	42.56	5.01	6.00	1,327,855	23,454.43	84.12	52.00	0.80
2			✓		✓		4.07	46.63	0.57	6.57	152,010	23,677.85	84.92	52.49	0.81
4	✓	✓		✓	✓		0.55	47.18	0.08	6.65	20,566	24,423.93	87.60	54.15	0.84
3	✓		✓	✓			2.44	49.62	0.34	7.00	91,206	24,572.74	88.13	54.48	0.84
3			✓	✓	✓		17.56	67.18	2.48	9.47	656,327	25,008.94	89.70	55.45	0.86
2	✓			✓			1.24	68.42	0.18	9.65	46,497	26,158.82	93.82	57.99	0.90
3	✓	✓		✓			0.12	68.54	0.02	9.67	4,471	26,281.87	94.26	58.27	0.90
3	✓			✓	✓		11.15	79.69	1.57	11.24	416,687	26,609.51	95.44	58.99	0.91
4	✓	✓	✓	✓			0.12	**79.81**	0.02	11.26	4,471	28,115.95	100.84	62.33	**0.96**
2			✓	✓			0.07	79.88	0.01	11.27	2,683	30,011.22	107.64	66.54	1.03
3	✓	✓	✓				0.02	79.90	0.00	11.27	894	31,141.96	111.69	69.04	1.07
4	✓	✓	✓		✓		0.05	79.95	0.01	11.28	1,788	34,798.76	124.81	77.15	1.19
2	✓		✓				0.65	80.60	0.09	11.37	24,143	35,523.33	127.41	78.76	1.22
3	✓	✓			✓		0.12	80.72	0.02	11.38	4,471	36,036.05	129.25	79.89	1.24
2	✓				✓		4.78	85.50	0.67	12.06	178,836	37,939.10	136.07	84.11	1.30
2				✓	✓		6.82	92.32	0.96	13.02	254,841	38,408.85	137.76	85.15	1.32
1			✓				0.22	**92.54**	0.03	13.05	8,048	43,180.52	154.87	95.73	**1.48**
1					✓		4.57	97.11	0.64	13.70	170,788	43,774.04	157.00	97.05	1.50
3		✓		✓	✓		0.19	97.30	0.03	13.72	7,153	48,074.64	172.42	106.58	1.65
1	✓						0.62	97.92	0.09	13.81	23,249	48,558.70	174.16	107.66	1.67
4		✓	✓	✓	✓		0.07	97.99	0.01	13.82	2,683	49,883.99	178.91	110.59	1.71
0							1.00	**99.00**	0.14	13.96	37,555	57,633.87	206.71	127.77	**1.98**
1				✓			0.84	99.83	0.12	14.08	31,296	63,869.23	229.07	141.60	2.19
2	✓	✓					0.17	100.00	0.02	14.10	6,259	68,692.05	246.37	152.29	2.36

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.48: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Development Region Level, Far-Western Region

							Far-Western Region								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Gender	**Population share of sample** (total obs.)*	**Cumulative Far-Western Region population share** (%)	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Far-Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~4~~	~~✓~~	~~✓~~	~~✓~~		~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
~~4~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~			~~0.00~~						~~0.00~~	~~0.00~~	
~~4~~		~~✓~~	~~✓~~	~~✓~~	~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
~~2~~		~~✓~~			~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~		~~✓~~	~~✓~~	~~✓~~			~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~	~~✓~~	~~✓~~		~~✓~~			~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~	~~✓~~	~~✓~~	~~✓~~				~~0.00~~						~~0.00~~	~~0.00~~	
~~2~~		~~✓~~		~~✓~~			~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~		~~✓~~		~~✓~~	~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
~~2~~	~~✓~~	~~✓~~					~~0.00~~						~~0.00~~	~~0.00~~	
~~1~~		~~✓~~					~~0.00~~						~~0.00~~	~~0.00~~	
~~3~~		~~✓~~	~~✓~~		~~✓~~		~~0.00~~						~~0.00~~	~~0.00~~	
~~2~~		~~✓~~	~~✓~~				~~0.00~~						~~0.00~~	~~0.00~~	
3	✓	✓			✓		0.15	0.15	0.01	0.01	3,577	14,673.33	50.85	32.53	0.50
5+Gender	✓	✓	✓	✓	✓	✓Female	0.22	-	0.02	-	5,365	14,679.09	50.87	32.54	0.50
5	✓	✓	✓	✓	✓		0.51	0.66	0.05	0.06	12,518	15,308.20	53.05	33.94	0.52
5+Gender	✓	✓	✓	✓	✓	✓Male	0.29	-	0.03	-	7,153	15,780.04	54.69	34.98	0.54
4	✓		✓	✓	✓		21.79	22.45	2.01	2.08	533,824	19,870.33	68.86	44.05	0.68
2	✓		✓				0.36	22.81	0.03	2.11	8,942	21,278.84	73.74	47.18	0.73
1	✓						0.77	23.58	0.07	2.18	18,778	24,195.23	83.85	53.64	0.83
3			✓	✓	✓		25.99	49.56	2.40	4.58	636,655	25,283.37	87.62	56.05	0.87
4	✓	✓		✓	✓		0.18	**49.74**	0.02	**4.60**	4,471	26,020.70	90.18	57.69	**0.89**
2	✓			✓			1.20	50.95	0.11	4.71	29,508	26,118.60	90.52	57.91	0.90
3	✓		✓	✓			1.50	52.45	0.14	4.85	36,661	27,199.72	94.26	60.30	0.93
3	✓		✓		✓		3.61	56.06	0.33	5.18	88,524	27,866.06	96.57	61.78	0.96
2			✓		✓		5.00	61.06	0.46	5.65	122,502	27,877.00	96.61	61.80	0.96
3	✓			✓	✓		8.47	**69.53**	0.78	6.43	207,449	27,954.03	96.88	61.97	**0.96**
2	✓				✓		5.00	74.53	0.46	6.89	122,502	31,477.72	109.09	69.79	1.08
2				✓	✓		13.03	87.55	1.20	8.10	319,222	34,740.37	120.39	77.02	1.19
1					✓		6.72	**94.27**	0.62	8.72	164,529	42,897.66	148.66	95.10	**1.47**
1			✓				0.51	94.78	0.05	8.76	12,518	43,999.33	152.48	97.55	1.51
0							2.04	96.82	0.19	8.95	50,074	46,910.37	162.57	104.00	1.61
2			✓	✓			1.75	**98.58**	0.16	**9.12**	42,921	56,224.20	194.85	124.65	**1.93**
1				✓			1.42	100.00	0.13	9.25	34,873	63,735.13	220.88	141.30	2.19

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.49: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Development Region Level, Eastern Region

											Eastern Region				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth Region (M/H/T)	Gender	**Population share of sample** (total obs.)*	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Eastern Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
5+Gender+M/H/T	✓	✓	✓	✓	✓	**M**	Male	0.81	0.19	0.19	49,180	28,467.29	70.91	63.11	0.98
5+Gender+M/H/T	✓	✓	✓	✓	✓	**M**	Female	1.03	0.24	0.42	62,592	30,325.48	75.54	67.23	1.04
5+Gender+M/H/T	✓	✓	✓	✓	✓	**H**	Female	3.17	0.73	1.15	193,143	30,606.27	76.24	67.85	1.05
5+Gender+M/H/T	✓	✓	✓	✓	✓	**H**	Male	2.80	0.64	1.80	170,788	30,645.13	76.33	67.94	1.05
5+Gender	✓	✓	✓	✓	✓		Female	5.33	1.23	-	325,481	31,230.42	77.79	69.24	1.07
5+Gender	✓	✓	✓	✓	✓		Male	4.47	1.03	-	272,724	31,979.77	79.66	70.90	1.10
5+Gender+M/H/T	✓	✓	✓	✓	✓	**T**	Female	1.14	0.26	2.06	69,746	33,770.97	84.12	74.87	1.16
5+Gender+M/H/T	✓	✓	✓	✓	✓	**T**	Male	0.86	0.20	2.26	52,757	39,574.72	98.57	87.74	1.36

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.50: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Development Region Level, Central Region

											Central Region				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth Region (M/H/T)	Gender	**Population share of sample** (total obs.)*	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Central Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Female	0.02	0.01	0.01	1,788	19,073.46	32.45	42.29	0.65
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Female	0.68	0.23	0.24	60,804	20,832.30	35.44	46.19	0.71
5+Gender	✓	✓	✓	✓	✓		Female	1.60	0.54	-	142,174	23,445.68	39.89	51.98	0.80
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Male	0.83	0.28	0.52	74,217	23,806.56	40.50	52.78	0.82
5+Gender	✓	✓	✓	✓	✓		Male	1.77	0.59	-	157,375	25,144.17	42.78	55.74	0.86
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Female	0.90	0.30	0.82	79,582	25,540.67	43.45	56.62	0.88
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Male	0.93	0.31	1.13	82,264	26,103.37	44.41	57.87	0.90
5+Gender+M/H/T	✓	✓	✓	✓	✓	**M**	**Male**	0.01	0.00	1.13	894	47,920.12	81.53	106.24	1.64

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.51: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Development Region Level, Western Region

								Western Region							
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth Region (M/H/T)	Gender	**Population share of sample** (total obs.)*	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~5+Gender+M/H/T~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~M~~	~~Male~~	~~0.00~~	-	-	-	-	~~0.00~~	~~0.00~~	-
~~5+Gender+M/H/T~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~M~~	~~Female~~	~~0.00~~	-	-	-	-	~~0.00~~	~~0.00~~	-
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Female	1.16	0.23	0.23	61,698	23,680.66	49.83	52.50	0.81
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Male	1.13	0.23	0.46	59,910	24,659.37	51.88	54.67	0.85
5+Gender	✓	✓	✓	✓	✓		Male	1.30	0.26	-	68,852	28,323.97	59.59	62.79	0.97
5+Gender	✓	✓	✓	✓	✓		Female	1.51	0.30	-	80,476	28,559.74	60.09	63.32	0.98
5+Gender+M/H/T	✓	✓	✓	✓	✓	**H**	**Female**	0.35	0.07	0.53	18,778	44,591.01	93.82	98.86	1.53
5+Gender+M/H/T	✓	✓	✓	✓	✓	**H**	**Male**	0.17	0.03	0.56	8,942	52,876.80	111.26	117.23	1.81

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.52: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Development Region Level, Mid-Western Region

								Mid-Western Region							
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth Region (M/H/T)	Gender	**Population share of sample** (total obs.)*	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Mid-Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~5+Gender+M/H/T~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~M~~	~~Male~~	~~0.00~~	-	-	-	-	~~0.00~~	~~0.00~~	-
~~5+Gender+M/H/T~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~M~~	~~Female~~	~~0.00~~	-	-	-	-	~~0.00~~	~~0.00~~	-
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Female	0.57	0.08	0.08	21,460	20,067.34	71.97	44.49	0.69
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Male	0.69	0.10	0.18	25,931	20,416.48	73.22	45.26	0.70
5+Gender	✓	✓	✓	✓	✓		Male	1.32	0.19	-	49,180	21,227.36	76.13	47.06	0.73
5+Gender	✓	✓	✓	✓	✓		Female	0.86	0.12	-	32,190	22,122.88	79.34	49.05	0.76
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Male	0.62	0.09	0.27	23,249	22,131.81	79.38	49.07	0.76
5+Gender+M/H/T	✓	✓	✓	✓	✓	T	Female	0.29	0.04	0.31	10,730	26,233.96	94.09	58.16	0.90

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.53: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Development Region Level, Far-Western Region

								Far-Western Region							
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Rural Birth Region	Birth Region (M/H/T)	Gender	**Population share of sample** (total obs.)*	**Estimated share overall sample population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of Far-Western Region mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~5+Gender+M/H/T~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~✓~~	~~H~~	~~Male~~	~~0.00~~	-	-	-	-	~~0.00~~	~~0.00~~	-
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Female	0.15	0.01	0.01	3,577	13,638.20	47.26	30.24	0.47
5+Gender+M/H/T	✓	✓	✓	✓	✓	M	Male	0.18	0.02	0.03	4,471	13,939.02	48.31	30.90	0.48
5+Gender+M/H/T	✓	✓	✓	✓	✓	H	Female	0.04	0.00	0.03	894	14,673.33	50.85	32.53	0.50
5+Gender	✓	✓	✓	✓	✓		Female	0.22	0.02	-	5,365	14,679.09	50.87	32.54	0.50
5+Gender	✓	✓	✓	✓	✓		Male	0.29	0.03	-	7,153	15,780.04	54.69	34.98	0.54
5+Gender+M/H/T	✓	✓	✓	✓	✓	**T**	**Female**	0.04	0.00	0.04	894	18,848.40	65.32	41.79	0.65
5+Gender+M/H/T	✓	✓	✓	✓	✓	**T**	**Male**	0.11	0.01	0.05	2,683	18,848.40	65.32	41.79	0.65

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *6,823 obs. in the eastern region, 9,943 obs. in the central region, 5,944 obs. in the western region, 4,180 obs. in the mid-western region, 2,740 obs. in the far-western region. **Total of 29,630 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

3.5.2.4. Robustness Checks: Economies of Scale, Bootstrapping, Multicollinearity, OLS Regression Validity Checks, etc.

Great care is taken to ensure the most coherent, clean and robust inequality of outcome and IOp analyses. To do so, at every stage of the analytical process, a series of measures was taken to check the data for consistency, the methodology for coherence and the statistical tools and results for robustness.

Refer to the national level robustness checks section (3.4.2.3) for the general robustness checks during the various stages for the analyses. These were also carried out at the Development Region level. The same categories are listed below. Only complementary and Development Region level specific remarks are given here. The remaining are similar to the national level analyses.

1. **Dataset**: e.g. data cleaning for outliers, missing values (especially Section 3.3.2)

2. **Dependant variable aggregations**: e.g. data cleaning, outliers, literature guidance comparison, unit conversion checks (Section 2.4 for consumption and Section 2.5 for income)

3. **Inequality in outcome indices:** e.g. various different indices computations with various dependant variables, bootstrapped standard errors (Sections 2.2 and 2.6)

4. **Type distribution and circumstance categorical checks and variations:** e.g. low type distribution checks, circumstance categorical variable combination checks (mainly Sections 3.1.4 and 3.3.2.5)

Additionally to the national level remarks, note that the sample populations for the mid- and far-west regions are relatively small (905 and 582, respectively). This lead to a relatively low level of type observation and a relatively large share of types with less than 5 observations.

5. **Regressions:** e.g. various regression tests and various regressions (e.g. Link test, Ramsey test, Stepwise regressions, residual checks for normality and homoscedasticity, multicollinearity check, variance inflation factor, outlier and leverage checks, Dfbeta test), (RF-OLS regression sections for each level of analyses: Section 3.4.2.3 for the National level analyses, Section 3.5.2.4 for the Development Region level, Section 3.6.2.4 for the urban rural level, Section 3.7.2.4 for the grouping my population quarter level)

The presented regression results in the relevant section are do not change drastically from one to another regression within each population sample. Most results are significant at the 1% level. R-squared is higher for the consumption than the income panel. Compared to the literature, this level of R-squared is seen as acceptable. In general, the consumption panel results are preferred over the income ones.

A series of regression tests were carried out to test for their validity, homoscedasticity, etc. These also hold for this level's analyses. Refer to the national level robustness section for more details.

6. **IOp indices:** e.g. parametric and non-parametric, scale equivalence checks, axiom satisfaction checks, income and consumption expenditure analyses comparisons (IOp scalar indices and IOp equivalence scale sections for each level of analyses)

Table 3.54 presents the equivalence scale robustness checks results for this level's scalar IOp analyses. For the consumption panel, the equivalence scale measures seem robust and consistent. In the east and centre total levels of inequalities increase with the equivalence parameters. For the other regions it marginally falls. This can possibly be due to a) the smaller population samples in these provinces, and b) much lower consumption variance (i.e. great welfare homogeneity).

When looking at the income panel, the initially computed IOL and IOR estimates for the mid-west are higher than the consumption panel ones (contradictory to all other regions, where consumption IOR is higher than income IOR) and the newly computed indices are extremely sensitive to scale measures. This can be due to the a) the relatively small sample size of mid-west, and b) a relatively large variance in income distribution in the region.

Overall, the equivalence scale measures further show that the IOp indices are robust. This is apart from the mid-west results in the income panel though.

7. **Results**: e.g. comparison between different level analyses of this study, and to other empirical results in the literature (corresponding results interpretations sections for all level analyses, at both the descriptive and econometric level)

Where striking, the Development Region level results are compared to the national level ones. Attempts to interpret particular results, such as the parental education ones, are done.

Table 3.54: Equivalence Scale Robustness Checks on the Scalar Inequality of Opportunity Indices, Development Region Level

	Eastern Region, all HH members *	**Equivalence Scale Application**		**Central Region,** all HH members **	**Equivalence Scale Application**		**Western Region,** all HH members ***	**Equivalence Scale Application**		**Mid-Western Region,** all HH members°	**Equivalence Scale Application**		**Far-Western Region,** all HH member °°	**Equivalence Scale Application**	
		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75
Per capita consumption expenditure (variable *pccons*)															
Total inequality (E0)	0.179	0.217	0.217	0.254	0.302	0.302	0.267	0.259	0.259	0.137	0.132	0.132	0.184	0.132	0.132
Non-parametric estimates															
IOL	0.064	0.066	0.066	0.112	0.161	0.161	0.105	0.081	0.081	0.039	0.030	0.030	0.071	0.030	0.030
IOR	0.357	0.306	0.306	0.441	0.535	0.535	0.393	0.314	0.314	0.285	0.232	0.232	0.388	0.232	0.232
Parametric estimates															
IOL	0.030	0.075	0.077	0.093	0.158	0.157	0.085	0.090	0.089	0.024	0.035	0.032	0.059	0.035	0.032
IOR	0.183	0.346	0.354	0.365	0.523	0.519	0.319	0.349	0.345	0.171	0.266	0.242	0.269	0.266	0.242
Per capita income (variable *pcincome*)															
Total inequality (E0)	0.484	0.541	0.541	0.490	0.567	0.567	0.573	0.564	0.564	0.632	**0.917**	**0.917**	0.837	0.736	0.736
Non-parametric estimates															
IOL	0.103	0.091	0.091	0.137	0.186	0.186	0.100	0.185	0.185	0.201	**0.480**	**0.480**	0.197	0.153	0.153
IOR	0.213	0.168	0.168	0.280	0.329	0.329	0.175	0.328	0.328	0.318	**0.523**	**0.523**	0.236	0.208	0.208
Parametric estimates															
IOL	0.050	0.103	0.116	0.092	0.173	0.176	0.067	0.083	0.085	0.072	**0.329**	**0.345**	0.077	0.120	0.117
IOR	0.103	0.190	0.215	0.188	0.305	0.310	0.116	0.147	0.151	0.114	**0.359**	**0.376**	0.091	0.163	0.158

Note: For reference see Buhmann et al. (1998, Coulter et al. (1992).

*For all 1,636 obs. (*pcons*) and 1,636 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

**For all 2,469 obs. (*pcons*) and 2,464 obs. (*pcincome*) obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

***For all 1,448 obs. (*pcons*) and 1,423 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°For all 899 obs. (*pcons*) and 894 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

°°For all 582 obs. (*pcons*) and 580 obs. (*pcincome*) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.6. Inequality of Opportunity Analysis in Nepal: Urban vs. Rural Areas

This section (3.6) focuses on the empirical analyses of IOp at the urban versus rural area level in Nepal. First, this section elaborates on the descriptive statistics.

Second, the empirical results are given and interpreted. To do so, initially the scalar indices of IOp, then the RF-OLS results and the partial IORs are given discussed. Following the circumstantial specific shares of unequal opportunities for the most opportunity deprived are highlighted and discussed. The descriptive and empirical findings of the urban rural area level analyse are combined and fully discussed.

Finally, various robustness checks are carried out and discussed.

3.6.1. Descriptive Analysis: Urban vs. Rural Areas

This section looks at some general noteworthy descriptive statistics that are different to the ones already elaborated and explained in the descriptive statistics section for the national and Development Region level analyses. It then discusses the type divisions and distributions for the urban and rural area level analyses. Finally, the descriptive statistics for circumstance variables and their sub-categories are given with an initial attempt of interpretation.

The descriptive statistics are given for a) urban and b) rural areas. All previously utilised circumstance variables are used apart from Birth Region 1 (urban rural area). Note that all HH members who fall into the 30 to 49 age range and have complete information on all circumstances are included into the analyses. This is because the urban rural area analyses is Nepal specific and one does not need to apply the HH head and spouse constraint for results comparability to the literature directly.

The results for per capita income are almost identical to those of per capita consumption and are thus not reported in the tables (available upon request). Only the descriptive type distribution table presents both income and consumption results.

The number of **observations for urban areas** is much smaller than for rural areas (880 versus 6,174; Table 3.57). This results in a relatively low sample **type distribution** for urban areas (Table 3.59). Out of 144 possible types, only 64 are observed in urban compared to 103 in rural areas.

Families in urban areas see a much lower presence of husbands and spouses and a relatively higher presence of sons, daughters, children-in-law and siblings-in-law compared to the rural areas. This goes in hand with the descriptive statistics of Section 3.5.1 for the Development Regions that suggest a) pull migration factors to the urban centres (and the central region), and b) the necessity to rely on family networks in order to deal with the higher costs of living in urban centres. This is further backed by the exceptionally high sex ratio in urban centres of over 100 (compared to 77 in rural areas). This suggests urban migration of particularly male labour migrants. The urban sex ratio level is much higher than the national and the central region average (90 and 89, respectively).

As expected, the proportion of **discriminated caste groups** is much higher in urban areas than compared to the national and rural average (80%, 70% and 66%, respectively). Combined with the HH composition and the sex ratio statistics, this suggests that especially individuals from discriminated caste groups tend to emigrate to urban centres. This is in contrast with the results of the IOp analyses, which suggests that actually, it is the most opportunity deprived individuals (i.e. particularly those who belong to the middle or lower castes) that face a disproportionate amounts of discrimination and opportunity deprivation in urban areas. So why emigrate to urban centre? An attempt for interpretation is given in the scalar IOp indices and RF-OLS regression sections that follow.

As expected, the level of **parental education** in the urban centres is more than 20% to 30% higher than in rural areas (Table 3.55). Mothers' education levels are still lower than those of fathers' (68% versus 43% of mothers have no schooling background, are illiterate or have not completed primary schooling), but yet much higher than in the rural areas (68% versus 89%, respectively).

On a side note, the descriptive statistics for **Birth Region 2** (Mountain, Hill, Terai) are consistent with the literature. Especially the mountain belt has little urban centres, as defined by the CBS (2011, Ch. 1).

Table 3.55: Definition and Description of Circumstance Variables, Urban Rural Level Analysis

	Grouping	**Expected Effects on y for RF-OLS Reg.**	**Urban Region,** Descriptive stats (%), all HH members*	**Rural Region,** Descriptive stats (%), all HH members**
Gender				
Category 1	Dominant (Male)	Positive	**50.11**	43.39
Category 2	Dominated (Female)	Negative	49.89	56.61
Ethnicity				
Category 1	Dominant caste/ethnic majority	Positive	20.23	33.92
Category 2	Dominated caste/ethnic minority	Negative	**79.77**	66.08
Religion				
Category 1	Hindu	Positive	85.23	84.06
Category 2	Buddhist	Positive	**10.23**	8.89
Category 3	Other	Negative	**4.55**	7.05
Father's education				
Category 1	No schooling illiterate primary incomplete	Negative	**43.52**	67.78
Category 2	Primary completed and above	Positive	56.48	32.22
Mother's education				
Category 1	No schooling illiterate primary incomplete	Negative	**68.18**	89.15
Category 2	Primary completed and above	Positive	31.82	10.85
Birth region 2 (M/H/T)				
Category 1	Mountain	Negative	**2.27**	10.37
Category 2	Hill	Positive	**68.98**	57.58
Category 3	Terai	Positive	**28.75**	32.05

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).
Note 2: Parental education was filled using the educational statistics for HHH/spouse brothers/sisters and sons/daughters.
Note 3: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).
*For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.56: Sex Ratio, Urban Rural Level Analyses

Sex ratio	
Urban Region, all HH members*	**Rural Region,** all HH members**
100.44	76.65

Note 1: Author's computations using the NLSS III dataset.
Note 2: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).

Table 3.57: Relationship to Household Head, Urban Rural Level Analyses (Variable *pccons*)

	Urban Region, all HH members*			**Rural Region,** all HH members**		
Relationship to HH head	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**
Household head	341	**38.75**	38.75	3,051	49.42	49.42
Husband/Wife	259	**29.43**	68.18	2,230	36.12	85.54
Son/Daughter	139	**15.8**	83.98	397	**6.43**	91.97
Brother/Sister	28	**3.18**	87.16	90	1.46	93.42
Father/Mother	1	0.11	87.27	37	0.6	94.02
Grandchild				4	0.06	94.09
Nephew/Niece	3	0.34	87.61	2	0.03	94.12
Son/Daughter-In-Law	76	**8.64**	96.25	296	**4.79**	98.91
Brother/Sister-In-Law	29	3.3	99.55	37	0.6	99.51
Father/Mother-In-Law				2	0.03	99.55
Other Family Relative	3	0.34	99.89	21	0.34	99.89
Servant/Servants Relatives				3	0.05	99.94
Other Non-Related	1	0.11	100	4	0.06	100
Total	**880**	**100**		**6,174**	**100**	

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).
Note 2: Parental education was filled using the educational statistics for HHH/spouse brothers/sisters and sons/daughters.
Note 3: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).
*For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.58: Birth Place Descriptive Statistics, Urban Rural Region Level

Birth place Development Region	**Descriptive stats (%),** all HH members (no data constraints 29,722 obs.)	**Urban Region (%),** all HH members*	**Rural Region (%),** all HH members**
Eastern	23.05	15.34	24.31
Central	33.52	**64.89**	30.90
Western	20.03	10.91	21.96
Mid-western	14.18	5.34	13.90
Far-western	9.22	**3.52**	8.92
Total	**100.00**	**100.00**	**100.00**

Note: Author's computations using the NLSS III dataset and per capita consumption expenditure (*pccons*). The results are similar for per capita income (*pcincome*, available upon request).
*For all 880 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 26,174 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.59: Type Division Description, Urban Rural Level Analyses

	Per capita consumption expenditure (*pccons*)		**Per capita income** (*pcincome*)	
	Urban Region, all HH members*	**Rural Region,** all HH members**	**Urban Region,** all HH members*	**Rural Region,** all HH members**
Maximum number of types	**144**	144	144	144
Number of types observed	**64**	103	64	103
Mean number of observations per type	13.70313	59.70	13.75	59.94
Min. number of observations per type	1	1	1	1
Max. number of observations per type	77	631	77	631
Proportion of types with fewer than 5 observations but more than 0.	20.14 (**29**)	25.00 (36)	20.14 (29)	25.00 (36)

Note: *For all 880 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 26,174 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.6.2. Econometric Analysis for Inequality of Opportunity, Results and Results' Interpretation: Urban vs. Rural Areas

This econometric analysis section consists first of the scalar indices of IOp. Second, it lays out and interprets the results of the RF-OLS regressions of the circumstance variables on economic outcome. It also looks closely at the circumstantial specific shares of unequal opportunities, the partial IORs. Third, the most opportunity deprived types are ranked by their average mean economic outcome and their profiles are interpreted. Last, the section highlights some of the robustness checks that were carried out throughout the analysis.

3.6.2.1. Parametric and Non-parametric Scalar Indices of Inequality of Opportunity

Table 3.60 presents the results for the scalar IOp indices for the urban rural area level. As expected, the total inequality in outcome (MLD or E_0) are higher for income than for consumption.

For **both, the income and consumption panel**, the urban IOR estimates are double the rural estimates. For consumption, opportunity deprivation makes up 34% to 44% of total consumption inequality in urban areas, while this is "only" between 17% to 21% in rural areas.

There are only three cases in this entire dissertation where the level of IOL(income) is lower than IOL(consumption) and that only to a minor extent and only for the parametric approach. Yet IOR(income) lower than IOR(consumption). Thus, the IOR estimates are still consistent. One example is the urban analysis at the urban rural area level. One part of the explanation is that total inequality in consumption is proportionately higher than compared to the rural area and its ratio to total income inequality in urban area is higher than the between E0(rural/consumption) and E0(rural/income) and the level of IOL(urban/consumption) higher than IOL(urban/income). This means that IOR(income) is a lot lower than IOR(consumption). FG (2011) have one observation also where IOL(income) is lower than IOL(consumption), Guatemala.

One exception where IOR(income) is higher than IOR(consumption) is for the mid-western region at the Development Region level analysis. Refer to Section 3.5.2.1. for more details.

The **statistical differences** between the **IOL versus IOR, parametric versus non-parametric estimates** were tested. Table 3.91 shows all the results using the paired t-test. In all cases the null can be rejected and the means of the IOp estimates, no matter whether IOL, IOR, parametric or non-parametric, are different from each other.

Table 3.95 reports the results of the paired t-test to test for **statistical difference** between all levels of the IOp analyses, i.e. national, Development Region, urban rural area, and by income/consumption population quarter level. The correlation matrix is available upon request. The coloured boxes within the table show the within level results statistical difference comparison. At the within urban rural area level all the computed IOp indices are statistically different between the urban and the rural areas.

Arguably, Nepal's limited economic development and having a centralised state means that there are more, albeit still few, opportunities in urban areas and especially the capital. However, this also means that it is easier to discriminate on circumstances. Also, the greater number of "opportunity privileged" individuals in urban centres are in fierce competition and must use all social capital to seize opportunities. In other words, the high population density in urban centres for relatively few economic opportunities means that these are reserved for only the most opportunity privileged people. Individuals who accumulate opportunity depriving circumstances (see following section) are left with disproportionately fewer opportunities.

Nepal's rural areas are still very underdeveloped and efforts to boost rural catch-up were significantly disrupted during the political and violent conflicts from 1996 to 2005 (UNDP and National Planning Commission, 2019). One may argue that everyone suffers more or less "equally" in the impoverished and left behind rural areas of Nepal. Yet, ironically, especially men belonging to the most opportunity deprived types (i.e. who accumulate several negatively impacting circumstances; more in the following section) tend to out-migrate to urban centres, or emigrate abroad. This in consequence, however, leaves more opportunities for the few people who remain in the rural areas,

also contributing to a relatively lower share of opportunity deprivation out of overall inequality in outcome.

Table 3.60: Scalar Indices of Inequality of Opportunity, Urban Rural Level Analyses

	Urban Region, all HH members*	**Rural Region,** all HH members**
Per capita consumption expenditure (variable *pccons*)		
Total inequality (E0)	0.275	**0.211**
Non-parametric estimates		
IOL	0.121	0.046
IOR	**0.442**	0.219
Parametric estimates		
IOL	0.094	0.036
IOR	**0.343**	0.170
Per capita income (variable *pcincome*)		
Total inequality (E0)	0.464	**0.550**
Non-parametric estimates		
IOL	0.151	0.071
IOR	**0.326**	0.128
Parametric estimates		
IOL	0.085	0.041
IOR	**0.184**	0.074

Note: *For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.6.2.2. Reduced-Form OLS Regressions and Circumstance Specific Shares of Opportunities (Partial IORs)

The RF-OLS regression results for the urban rural level analyses on per capita consumption expenditure are reported in Table 3.61 and on per capita income in Table 3.62. For comparison purpose to the national level and Development Region level analyses, the same regression and categorical variable combinations are run and presented here.

Table 3.65 shows the ***partial IOR*** estimates, that is the circumstance specific shares of overall opportunities compared to overall inequality in outcome.

In general, whatever the effect of a circumstance variable (positive or negative), the impact of it in urban areas is almost always larger than in rural areas for all regressions and the partial IOR results. In other words, having characteristics that are doomed to hinder opportunities (or enhance them) are seen to be more impacting in urban than rural areas. For instance, the degree of caste or religious minority discrimination in the urban areas is higher than in rural areas. Suggesting that if one accumulates several negative circumstances, then one should remain in the rural areas, and the more positive circumstances one accumulates, one is encouraged to move to the urban centres.

Upfront, it appears important to highlight that the differences in amplitudes of the regression coefficients between the consumption and income panels are interesting. The divergences later elaborated go in hand with the literature review on the measure of welfare and economic well-being elaborated in Chapter 2.

For both urban economic welfare panels, **mother's education** appears to be much more important than **father's**. Especially when looking at the income regressions, mother's education dominates over father's. For the rural regressions, father's education levels are more important for the consumption panel, while mother's are more important for the income panel. These regression results are confirmed by the partial IOR estimates, where mother's education explains over 25% and father's education only 21% of overall consumption inequality, and the effect amplitude of parental education on opportunities is a lot higher in urban than rural areas. In rural areas, father's education is by 4% points more important than mother's. The results between the rural and urban areas, between mother's and father's education and between the OLS and the partial IOR results are all statistically different (paired t-test).

At first, these results may appear contradictory to those revealed by the Development Region analyses, but they are not. It is important to bear in mind that the division of the Development Regions is done for administrative purpose, and that these regions include both urban and rural areas. Thus, the Development Region coefficients are both, urban and rural averaged estimates for each region. As can be seen from the descriptive statistics, more than 85% of Nepal's population lives in rural areas. So when estimates are run at the Development Region level, the regression samples are dominated by the rural populations. The urban rural analyses, however, completely segregate these two populations and it becomes evident that mother's education clearly dominates over father's for children's future economic outcome. Yet, in rural areas and for consumption welfare, father's education is almost twice as important as mother's. In rural areas for the income panel, father's and mother's education impact more or less equally on children's future income levels.

The urban rural level regression analyses results are consistent with the literature on the importance of maternal education attainment for offspring.

The partial IORs suggest that in urban areas, belonging to an **ethnic minority** or a lower caste group are important circumstances in explaining opportunity differentials explaining over 34% of overall consumption ienquality, while **religious minority** group belonging (Buddhist or other religious minorities) accounts for less than 8% each. Being born in the hill region and urban areas is estimated to positively enhance opportunities by explaining over 10% of overall consumption inequality. This is consistent with the explanations given above. The rural partial IORs are a lot more moderate compared to those of the urban area, one main explanations being that the overall parametric IOR (and MLD) estimates is lower.

The share of ethnic and caste group belonging in the partial IOR estimates for urban areas accounting for over 34% of outcome inequality is the highest value taken on by any partial IOR across all four levels of analyses in this dissertation. Jodhka and Shah (2010, p. 102) find in their extensive study of residential discrimination of the Dalit community (i.e. the untouchables or lowest ranking caste groups) in Nepal and other South Asian countries that still today they suffer "from physical touch and residential segregation to taboos and restrictions on inter-dining (… and) physical movement (…). They seem to be their living in segregated settlements away from the main village, or in the urban slums where living conditions are generally poor." Yet, they also highlight that the discrimination against the Dalit community has been declining in Nepal over the past two decades.

The comparison between the urban and the rural partial IORs confirms the RF-OLS results where a) the effect (both positive and negative) of circumstance weigh significantly more in urban than rural areas, b) the parental education level is an important factor explaining differences in outcome and combined accounting for up over 46% of opportunities, and c) the level of mother's education is statistically different and more important than that of father's in urban areas (reverse in rural areas).

The statistical difference between parental education and between the urban and rural areas (both RF-OLS and partial IORs) were tested to be statistically different from each other.

Despite the relatively low sample population and subsequent type distribution, the IOp and RF-OLS regression analyses are robust, with high levels of statistical significance, for most independent variables. Also, **R-squared** for urban areas is amongst the highest in all regression analyses in this study and higher than for all regression results for all 6 Latin American countries analyses by FG (2011).

Note that the correlation matrices here are only for per capita consumption. The ones for income are available upon request. Nevertheless, when comparing the correlation matrices for consumption, it appears that belonging to a dominated caste group negatively correlates with parental education in both urban and rural areas.

Table 3.61: Reduced-Form OLS Regressions of Circumstances on Per Capita Consumption Expenditure, Urban Rural Level

Circumstances	**Urban Region,** for all HH members				**Rural Region,** for all HH members			
	ln(pcconso)	ln(pcconso)	ln(pcconso)	ln(pcconso)	ln(pcconso)	ln(pcconso)	ln(pcconso)	ln(pcconso)
Ethnic Minority / Dominated Caste	-0.189***	-0.189***	-0.249***	**-0.251*** **	-0.213***	-0.214***	-0.216***	**-0.217*** **
	(0.059)	(0.059)	(0.061)	(0.061)	(0.020)	(0.020)	(0.019)	(0.019)
Buddhist	0.245***	0.245***	0.457***	**0.457*** **	0.147***	0.147***	0.146***	0.145***
	(0.054)	(0.054)	(0.057)	(0.057)	(0.032)	(0.032)	(0.031)	(0.031)
Other Religion	-0.165	-0.164	-0.433***	**-0.436*** **	-0.111***	-0.110***	-0.112***	-0.112***
	(0.111)	(0.111)	(0.124)	(0.122)	(0.034)	(0.034)	(0.033)	(0.033)
Father's Education: Primary Complete or Above	0.237***	0.237***	0.344***	**0.344*** **	0.360***	0.361***	0.363***	**0.364*** **
	(0.052)	(0.052)	(0.055)	(0.055)	(0.020)	(0.020)	(0.020)	(0.020)
Mother's Education: Primary Complete or Above	0.374***	0.374***	0.405***	**0.404*** **	0.192***	0.189***	0.193***	**0.189*** **
	(0.054)	(0.054)	(0.058)	(0.058)	(0.028)	(0.028)	(0.028)	(0.028)
Hill Birth Region	0.361*	0.360*			0.122***	0.122***		
	(0.186)	(0.185)			(0.028)	(0.028)		
Terai Birth Region	-0.269	-0.269			0.083***	0.082***		
	(0.185)	(0.185)			(0.029)	(0.029)		
Female	-0.004		0.022		0.016		0.017	
	(0.045)		(0.051)		(0.016)		(0.016)	
Constant	10.751***	10.749***	10.826***	10.839***	10.209***	10.219***	10.304***	10.314***
	(0.192)	(0.190)	(0.074)	(0.064)	(0.031)	(0.030)	(0.020)	(0.018)
Observations	877	877	877	877	6149	6149	6149	6149
R-squared	0.407	0.407	0.267	0.267	0.149	0.149	0.145	0.145
R-squared (adjusted)	0.401	0.402	0.262	0.263	0.148	0.148	0.145	0.145
F stat	80.154	91.403	48.293	58.111	120.539	137.549	153.499	183.831

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.62: Reduced-Form OLS Regressions of Circumstances on Per Capita Income, Urban Rural Level

Circumstances	**Urban Region,** for all HH members				**Rural Region,** for all HH members			
	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)	ln(pcincome)
Ethnic Minority / Dominated Caste	**-0.291*****	-0.284***	-0.340***	-0.335***	-0.162***	-0.162***	-0.159***	-0.158***
	(0.091)	(0.091)	(0.090)	(0.089)	(0.037)	(0.037)	(0.035)	(0.035)
Buddhist	**0.428*****	0.427***	0.589***	0.587***	0.232***	0.233***	0.225***	0.226***
	(0.090)	(0.091)	(0.087)	(0.088)	(0.053)	(0.053)	(0.051)	(0.051)
Other Religion	-0.287	-0.28	-0.477***	-0.471***	-0.247***	-0.247***	-0.250***	-0.250***
	(0.177)	(0.178)	(0.180)	(0.180)	(0.052)	(0.052)	(0.052)	(0.052)
Father's Education: Primary Complete or Above	**0.291*****	0.292***	0.379***	0.379***	**0.322*****	0.320***	0.325***	0.323***
	(0.084)	(0.084)	(0.083)	(0.083)	(0.035)	(0.035)	(0.035)	(0.035)
Mother's Education: Primary Complete or Above	**0.405*****	0.407***	0.435***	0.436***	**0.347*****	0.357***	0.349***	0.359***
	(0.084)	(0.084)	(0.086)	(0.086)	(0.055)	(0.055)	(0.055)	(0.055)
Hill Birth Region	**0.715*****	0.701***			0.108*	0.108*		
	(0.226)	(0.231)			(0.057)	(0.057)		
Terai Birth Region	0.256	0.247			0.086	0.088		
	(0.229)	(0.234)			(0.061)	(0.061)		
Female	-0.09		-0.066		-0.045		-0.044	
	(0.072)		(0.075)		(0.029)		(0.029)	
Constant	11.089***	11.047***	11.581***	11.542***	10.779***	10.753***	10.865***	10.839***
	(0.238)	(0.239)	(0.108)	(0.096)	(0.063)	(0.060)	(0.036)	(0.032)
Observations	877	877	877	877	6118	6118	6118	6118
R-squared	0.241	0.239	0.196	0.195	0.059	0.059	0.058	0.058
R-squared (adjusted)	0.234	0.233	0.19	0.19	0.058	0.057	0.057	0.057
F stat	27.51	31.297	27.476	32.86	39.846	45.234	52	62

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.63: Correlation Matric of Circumstances for All Household Members Using Per Capita Consumption Expenditure, Urban Rural Level Analyses, **Urban Area**

	Female	Ethnic Minority / Dominated Caste	Buddhist	Other Religion	Father's Education Incomplete or Above	Mother's Education Incomplete or Above	Hill Birth Region	Terai Birth Region
Female	**1**							
Ethnic Minority / Dominated Caste	**-0.0571** (0.0913)	**1**						
Buddhist	**0.017** (0.6160)	**0.1594** (0.0000)	**1**					
Other Religion	**-0.0315** (0.3508)	**0.0681** (0.0437)	**-0.0739** (0.0286)	**1**				
Father's Education Incomplete or Above	**0.0341** (0.3136)	**-0.1175** (0.0005)	**0.0629** (0.0626)	**-0.083** (0.0140)	**1**			
Mother's Education Incomplete or Above	**0.021** (0.5337)	**-0.1027** (0.0023)	**0.0289** (0.3929)	**-0.0898** (0.0078)	**0.3906** (0.0000)	**1**		
Hill Birth Region	**0.0479** (0.1565)	**-0.0522** (0.1222)	**0.1705** (0.0000)	**-0.2307** (0.0000)	**0.1776** (0.0000)	**0.0913** (0.0068)	**1**	
Terai Birth Region	**-0.0291** (0.3896)	**0.0408** (0.2270)	**-0.1656** (0.0000)	**0.2348** (0.0000)	**-0.1497** (0.0000)	**-0.0591** (0.0804)	**-0.9471** (0.0000)	**1**

Note: Display all pairwise correlation coefficients. Significance level for each entry is in parenthesis. Data sample constraints: For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Correlation matrices for per capita income are not displayed here, but available upon request.

Table 3.64: Correlation Matric of Circumstances for All Household Members Using Per Capita Consumption Expenditure, Urban Rural Level Analyses, **Rural Area**

	Female	Ethnic Minority / Dominated Caste	Buddhist	Other Religion	Father's Education Incomplete or Above	Mother's Education Incomplete or Above	Hill Birth Region	Terai Birth Region
Female	**1**							
Ethnic Minority / Dominated Caste	**-0.0184** (0.1491)	**1**						
Buddhist	**-0.014** (0.2738)	**0.2189** (0.0000)	**1**					
Other Religion	**0.002** (0.8783)	**0.1742** (0.0000)	**-0.086** (0.0000)	**1**				
Father's Education Incomplete or Above	**0.0153** (0.2304)	**-0.2011** (0.0000)	**-0.0144** (0.2586)	**-0.0261** (0.0410)	**1**			
Mother's Education Incomplete or Above	**-0.0899** (0.0000)	**-0.1104** (0.0000)	**-0.0153** (0.2300)	**-0.0408** (0.0014)	**0.2897** (0.0000)	**1**		
Hill Birth Region	**0.0326** (0.0106)	**-0.2171** (0.0000)	**0.0463** (0.0003)	**-0.0068** (0.5945)	**0.0609** (0.0000)	**-0.0159** (0.2135)	**1**	
Terai Birth Region	**-0.0481** (0.0002)	**0.279** (0.0000)	**-0.1485** (0.0000)	**0.0057** (0.6533)	**-0.0488** (0.0001)	**0.0316** (0.0132)	**-0.8001** (0.0000)	**1**

Note: Display all pairwise correlation coefficients. Significance level for each entry is in parenthesis.
Data sample constraints: For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
Correlation matrices for per capita income are not displayed here, but available upon request.

Table 3.65: Circumstantial Specific Opportunity Share (Partial IORs), Urban Rural Level

	Urban Region (all HH members)^	**Rural Region** (all HH members)^^
Per capita consumption expenditure (variable *pccons*)		
Mean Logarithmic Deviation	0.275	0.211
IOp Index Ratio (parametric)	0.343	0.170
Circumstances		
Ethnic Minority / Dominated Caste	**0.343**	0.033
Buddhist	0.085	0.034
Other Religion	0.061	0.078
Father's Education: Primary Complete or Above	**0.206**	0.076
Mother's Education: Primary Complete or Above	**0.248**	0.040
Rural Birth Region	-	-
Hill Birth Region	0.105	0.029
Terai Birth Region	-	0.027
Female	-	-
Per capita income (variable *pcincome*)		
Mean Logarithmic Deviation	0.464	0.550
IOp Index Ratio (parametric)	0.184	0.074
Circumstances		
Ethnic Minority / Dominated Caste	0.035	0.033
Buddhist	0.025	0.018
Other Religion	0.001	0.016
Father's Education: Primary Complete or Above	0.060	0.060
Mother's Education: Primary Complete or Above	0.077	0.040
Rural Birth Region	-	-
Hill Birth Region	0.003	0.033
Terai Birth Region	-	-
Female	-	-

Note: only circumstances that were statistically significant at least at the 10 percent level were used to compute the partial IORs.
^For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
^^For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.6.2.3. Most Opportunity Deprived Types and Their Profiles

This section further builds on the national and Development Region level results that rank the mean economic advantage of types of individuals. This section does the same but for urban rural area samples. The tables presented here are for both consumption and income. Table 3.68 through to Table 3.71 ranks the different types of populations by their mean per capita consumption for urban and rural areas for both income and consumption. Table 3.72 through to Table 3.75 dig further into the characteristics of the most opportunity deprived by further disaggregating them by gender and Birth Region 2 (Mountain, Hill, Terai).

For the sake of completeness, this section provides summary tables for the mean levels of per capita consumption and income for each area (Table 3.66), and the descriptive statistics for the type distribution (Table 3.67). Both tables use all observations in the NLSS III dataset that have complete information on circumstances and on consumption and income, respectively. The HH composition and age constraints are not applied. This is to render the ranking as nationally representative as possible. Almost all observations in the dataset are used.

There are some **circumstantial accumulation outliers** due to very low type distributions particularly for the urban area rankings. The urban area, for instance, observes types with only two

negatively impacting circumstantial variables (religious minority other than Buddhist and mother's education level) with a mean consumption level of below 44% of the urban mean.

In general, for income and consumption expenditure, the more negatively impacting circumstances one accumulates, the lower the mean economic outcome and the more likely one is to suffer from opportunity deprivation. That said, there is a significant difference between rural and urban areas in terms of the population share and the proportion of their economic welfare. That is, their sample sizes are vastly different. For instance, 37% of the urban population have consumption levels below the national average. This, however, represents only 5% of the national population. For rural areas, 79% of the rural populations have consumption levels below the national average, but this represents over 67% of the national population.

3% of the urban and 5% of the rural population belong to the most opportunity deprived. However, due to the rural population being much larger, this translates into over 1.1 million people in Nepalese countryside and "only" 125,000 people in Nepal's cities.

For consumption, over 94% of the rural and barely 37% of the urban population lives on less than 2USD/day. In absolute terms, this is over 21.1 million people in the countryside and 1.5 million people in the cities. Yet, in urban centres almost no one lives on below 1USD/day, while in the countryside this mounts to over 6% (i.e. about 140,000 people) of the rural population.

For consumption and the most opportunity deprived types who accumulate four or more negatively impacting circumstances, everyone is badly off. However, those individuals living in the urban regions of the hill area (which hosts the capital and other economic urban centres), are a little better off with their mean type consumption level above the national mean. Women who accumulate four negative circumstances there earn 135% of the national per capita mean consumption, yet only 80% of the urban mean. For the most opportunity deprived men in the hill region, they earn even 188% of the national mean consumption (or 112% of the urban mean). This suggests a) the most opportunity deprived should try to be in the hill area, and b) there appears to be a large gender gap with men earning 40% more than women.

For the most opportunity deprived types in rural areas, they are all badly off. While men and women in the rural hill areas are the least bad off amongst the most deprived, their per capita consumption is still below 69% or 65% of the national mean, and 78% or 74% of the rural mean for men and women, respectively.

Looking at the most opportunity deprived tables for income (Table 3.74 and Table 3.75), one can see both a gender gap and an ecological belt trend for the urban region. Women born in the mountain and Terai belts are the worst off amongst all of the most disadvantaged type groups for income. This is consistent with the literature that a) suggests that there is great gender inequality in Nepal, particularly in terms of labour market access, and b) that women carry out a lot of unproductive and non-remunerated work. The gender gap amongst the most disadvantaged disappears for people born in the hill region. People born in the urban areas of the hill region are the least worse off amongst the most disadvantaged. However, caution must be taken when interpreting this result, because the sample size for these types in urban areas is extremely small and they may represent outliers, because the results indicate that they earn about 211% and 230% of the national mean income, (i.e. 134% and 146% of the urban level income).

Table 3.66: Urban Rural Level Mean Per Capita Consumption and Income

		NRs	USD*	USD/day	Obs.
Mean per capita income		**99,612.93**	**1,246.72**	**3.42**	**29,676**
	Urban	157,068.11	1,965.81	5.39	4,527
	Rural	89,270.59	1,117.28	3.06	25,149
Mean per capita consumption expenditure		**45,129.35**	**564.82**	**1.55**	**29,588**
	Urban	75,941.57	950.46	2.60	4,522
	Rural	39,570.71	495.25	1.36	25,066

Note: For the author's purposed reference variables: *pccons* and *pcincome*.
*av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.67: Sample Partition Description, Urban Rural Level

	For Per Capita Consumption Expenditure *(pccons)**		**For Per Capita Income** *(pcincome)***	
	Urban area, (all HH members, without data constraints, 4,522 obs.)	**Rural area**, (all HH members, without data constraints, 25,066 obs.)	**Urban area**, (all HH members, without data constraints, 4,527 obs.)	**Rural area**, (all HH members, without data constraints, 25,149 obs.)
Maximum number of types	144	144	144	144
Number of types observed	102	119	102	119
Mean number of observations per type	44.33	210.64	44.38	211.34
Min. number of observations per type	1	1	1	1
Max. number of observations per type	251	2100	251	2101
Proportion of types with fewer than 5 observations but more than 0.	18.75 (27)	14.58 (21)	18.75 (27)	14.58 (21)

Note: *without data constraints, 29,588 obs.
**without data constraints, 29,676 obs.)

Table 3.68: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Urban Rural Area Level, **Urban Area**

							Urban Area								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative urban area population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of urban area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
2		✓	✓				0.00	.	.	.	.	.	.	0.00	.
4+Gender	✓	✓	✓	✓		✓Male	1.61	-	0.25	-	65,368	31,471.07	41.44	69.74	1.08
4	✓	✓	✓	✓			3.07	3.07	0.47	0.47	**124,467**	33,070.59	43.55	73.28	1.13
4+Gender	✓	✓	✓	✓		✓Female	1.46	-	0.22	-	59,100	34,839.76	45.88	77.20	1.19
2		✓		✓			0.24	3.32	0.04	0.51	9,850	39,602.23	52.15	87.75	1.36
3	✓	✓		✓			1.81	5.13	0.28	0.78	73,427	42,162.11	55.52	93.43	1.45
3		✓	✓	✓			0.02	5.15	0.00	0.79	895	43,068.45	56.71	95.43	1.48
3	✓	✓	✓				0.55	5.71	0.08	0.87	22,386	45,537.64	59.96	100.90	1.56
3	✓		✓	✓			22.03	27.73	3.37	4.24	891,866	54,591.32	71.89	120.97	1.87
2	✓		✓				7.36	35.10	1.13	5.36	298,184	56,836.44	74.84	125.94	1.95
2	✓	✓					1.64	**36.73**	0.25	**5.61**	66,263	57,156.69	75.26	126.65	**1.96**
2	✓			✓			17.03	53.76	2.60	8.22	689,495	70,513.24	92.85	156.25	2.42
1			✓				3.30	57.05	0.50	8.72	133,422	72,939.34	96.05	161.62	2.50
2			✓	✓			4.20	61.26	0.64	9.36	170,135	73,681.05	97.02	163.27	2.53
1		✓					0.22	61.48	0.03	9.40	8,954	93,410.92	123.00	206.98	3.20
1				✓			7.05	68.53	1.08	10.47	285,648	94,078.17	123.88	208.46	3.23
1	✓						18.93	87.46	2.89	13.37	766,503	99,250.08	130.69	219.92	3.40
0							12.54	100.00	1.92	15.28	507,719	107,833.90	142.00	238.94	3.70

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,522 obs. in the urban areas and 25,066 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.69: Most Opportunity Deprived Types and Their Profile for Per Capita Consumption Expenditure, Urban Rural Area Level, **Rural Area**

							Rural Area								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative rural area population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of rural area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender	✓	✓	✓	✓		✓Female	2.46	-	2.08	-	551,596	28,291.01	71.49	62.69	0.97
4	✓	✓	✓	✓			4.90	4.90	4.15	4.15	1,098,714	28,508.93	72.05	63.17	0.98
4+Gender	✓	✓	✓	✓		✓Male	2.44	-	2.07	-	547,118	28,728.63	72.60	63.66	0.99
3	✓	✓	✓				0.61	**5.51**	0.52	**4.67**	**137,899**	28,833.37	72.87	63.89	**0.99**
3	✓		✓	✓			38.04	43.55	32.22	36.89	8,537,198	31,519.23	79.65	69.84	1.08
2	✓		✓				4.92	48.46	4.17	41.06	1,104,087	32,200.77	81.38	71.35	1.10
3	✓	✓		✓			1.65	50.11	1.40	42.45	369,820	34,392.43	86.91	76.21	1.18
3		✓	✓	✓			0.07	50.18	0.06	42.51	16,118	35,785.83	90.44	79.30	1.23
1		✓					0.03	50.22	0.03	42.54	7,164	37,341.32	94.37	82.74	1.28
2			✓	✓			13.69	63.91	11.60	54.14	3,073,176	39,634.60	100.16	87.82	1.36
1			✓				2.77	66.68	2.35	56.49	622,336	39,921.67	100.89	88.46	1.37
2	✓			✓			12.17	78.85	10.31	66.80	2,732,011	42,479.35	107.35	94.13	1.46
2	✓	✓					0.40	79.25	0.34	67.14	90,440	50,608.88	127.89	112.14	1.74
2		✓		✓			0.09	79.34	0.07	67.22	19,700	51,391.91	129.87	113.88	1.76
1	✓						6.41	85.75	5.43	72.64	1,438,089	52,917.20	133.73	117.26	1.81
1				✓			8.21	**93.96**	6.96	**79.60**	**1,842,831**	53,243.78	134.55	117.98	**1.83**
0							6.04	100.00	5.11	84.71	1,354,812	67,979.23	171.79	150.63	2.33
2		✓	✓				0.00	100.00	0.00	84.72	895	72,397.82	182.96	160.42	2.48

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,522 obs. in the urban areas and 25,066 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.70: Most Opportunity Deprived Types and Their Profile for Per Capita Income, Urban Rural Area Level, **Urban Area**

							Urban Area								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative urban area population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of urban area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
~~2~~		~~✓~~	~~✓~~				~~0.00~~							~~0.00~~	
2		✓		✓			0.24	0.24	0.04	0.04	64,378	60,813.05	38.72	61.05	2.09
4+Gender	✓	✓	✓	✓		✓Male	1.61	-	0.25	-	427,236	64,816.18	41.27	65.07	2.22
3		✓	✓	✓			0.02	0.27	0.00	0.04	5,853	68,774.60	43.79	69.04	2.36
4	✓	✓	✓	✓			3.07	3.34	0.47	0.51	813,505	79,916.31	50.88	80.23	2.74
3	✓	✓	✓				0.55	3.89	0.08	0.59	146,314	84,093.27	53.54	84.42	2.88
3	✓	✓		✓			1.81	5.70	0.28	0.87	479,909	96,131.29	61.20	96.50	3.30
4+Gender	✓	✓	✓	✓		✓Female	1.46	-	0.22	-	386,268	96,617.98	61.51	96.99	3.31
2	✓	✓					1.63	7.33	0.25	1.12	433,089	98,530.88	62.73	98.91	3.38
3	✓		✓	✓			22.00	29.34	3.36	4.47	5,829,142	106,619.60	67.88	107.03	3.66
2	✓		✓				7.36	36.69	1.12	5.60	1,948,900	130,700.40	83.21	131.21	4.48
2			✓	✓			4.22	40.91	0.64	6.24	1,117,837	151,748.80	96.61	152.34	5.20
1		✓					0.22	41.13	0.03	6.27	58,526	153,936.10	98.01	154.53	5.28
2	✓			✓			17.01	**58.14**	2.59	**8.87**	4,506,465	154,514.50	**98.37**	155.11	5.30
1			✓				3.34	61.48	0.51	9.38	883,735	157,622.60	100.35	158.24	5.40
1				✓			7.07	68.54	1.08	10.46	1,872,817	197,683.50	125.86	198.45	6.78
1	✓						18.91	87.45	2.88	13.34	5,009,785	204,514.60	130.21	205.31	7.01
0							12.55	100.00	1.91	15.25	3,324,250	212,297.10	135.16	213.12	7.28

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,527 obs. in the urban areas and 25,149 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on income. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.71: Most Opportunity Deprived Types and Their Profile for Per Capita Income, Urban Rural Area Level, **Rural Area**

							Rural Area								
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative rural area population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of rural area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender	✓	✓	✓	✓		✓Male	2.47	-	2.09	-	70,308	52,417.64	58.72	52.62	1.80
4	✓	✓	✓	✓			4.93	**4.93**	4.18	4.18	**1,306,342**	53,924.82	60.41	54.13	**1.85**
4+Gender	✓	✓	✓	✓		✓Female	2.46	-	2.09	-	52,731	55,436.88	62.10	55.65	1.90
3	✓	✓		✓			1.64	6.57	1.39	5.57	435,096	60,512.55	67.79	60.75	2.07
3	✓	✓	✓				0.61	7.19	0.52	6.09	162,239	62,075.29	69.54	62.32	2.13
2		✓		✓			0.09	7.27	0.07	6.16	23,177	62,075.29	69.54	62.32	2.13
2	✓		✓				4.91	12.18	4.16	10.32	1,300,021	71,755.55	80.38	72.03	2.46
3	✓		✓	✓			37.98	50.16	32.19	42.51	10,063,044	74,359.60	83.30	74.65	2.55
2	✓	✓					0.40	50.56	0.34	42.85	106,404	78,258.61	87.66	78.56	2.68
3		✓	✓	✓			0.07	50.63	0.06	42.91	18,963	80,717.56	90.42	81.03	2.77
1			✓				2.78	53.42	2.36	45.27	737,451	86,793.35	97.23	87.13	2.98
2			✓	✓			13.72	**67.14**	11.63	**56.89**	3,634,579	87,777.70	98.33	88.12	3.01
2	✓			✓			12.16	79.30	10.31	67.20	3,222,660	89,794.06	100.59	90.14	3.08
1		✓					0.03	79.33	0.03	67.23	8,428	113,304.40	126.92	113.74	3.89
1				✓			8.25	87.58	6.99	74.22	2,184,962	120,747.50	135.26	121.22	4.14
2		✓	✓				0.00	87.58	0.00	74.22	1,054	124,113.70	139.03	124.60	4.26
1	✓						6.39	93.98	5.42	79.64	1,694,030	126,986.60	142.25	127.48	4.35
0							6.02	100.00	5.11	84.75	1,596,054	157,968.80	176.96	158.58	5.42

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,527 obs. in the urban areas and 25,149 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on income. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.72: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Urban Rural Area Level, **Urban Region**

											Urban Area				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative urban population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of urban area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender+M/H/T	✓	✓	✓	✓	T	Female	1.06	1.06	0.16	0.16	42,981	25,657.62	33.79	56.85	0.88
4+Gender+M/H/T	✓	✓	✓	✓	T	Male	1.42	2.48	0.22	0.38	57,309	25,699.48	33.84	56.95	0.88
4+Gender	✓	✓	✓	✓		Male	1.61	-	0.25	-	65,368	31,471.07	41.44	69.74	1.08
4+Gender+M/H/T	✓	✓	✓	✓	M	Female	0.02	2.50	0.00	0.38	895	33,357.76	43.93	73.92	1.14
4+Gender	✓	✓	✓	✓		Female	1.46	-	0.22	-	59,100	34,839.76	45.88	77.20	1.19
4+Gender+M/H/T	✓	✓	✓	✓	M	Male	0.07	2.57	0.01	0.39	2,686	47,686.53	62.79	105.67	1.64
4+Gender+M/H/T	✓	✓	✓	✓	H	Female	0.38	2.94	0.06	0.45	15,223	60,852.95	80.13	134.84	2.09
4+Gender+M/H/T	✓	✓	✓	✓	H	Male	0.13	3.07	0.02	0.47	5,373	84,927.03	111.83	188.19	2.91

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,522 obs. in the urban areas and 25,066 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.73: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Consumption Expenditure, Urban Rural Area Level, **Rural Region**

											Rural Area				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative rural population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of rural area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender+M/H/T	✓	✓	✓	✓	M	Male	0.24	0.24	0.20	0.20	53,727	23,615.48	59.68	52.33	0.81
4+Gender+M/H/T	✓	✓	✓	✓	T	Female	0.94	1.18	0.80	1.00	249,450	26,897.63	67.97	59.60	0.92
4+Gender+M/H/T	✓	✓	✓	✓	T	Male	1.00	2.18	0.84	1.85	264,247	27,110.33	68.51	60.07	0.93
4+Gender	✓	✓	✓	✓		Female	2.46	-	2.08	-	651,106	28,291.01	71.49	62.69	0.97
4+Gender	✓	✓	✓	✓		Male	2.44	-	2.07	-	645,821	28,728.63	72.60	63.66	0.99
4+Gender+M/H/T	✓	✓	✓	✓	M	Female	0.34	2.51	0.28	2.13	88,787	29,057.97	73.43	64.39	1.00
4+Gender+M/H/T	✓	✓	✓	✓	H	Female	1.18	3.69	1.00	3.13	312,869	29,184.30	73.75	64.67	1.00
4+Gender+M/H/T	✓	✓	✓	✓	H	Male	1.20	4.90	1.02	4.15	318,154	31,091.96	78.57	68.90	1.07

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,522 obs. in the urban areas and 25,066 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on consumption expenditure. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.74: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Income, Urban Rural Area Level, **Urban Region**

											Urban Area				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative urban population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of urban area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender+M/H/T	✓	✓	✓	✓	M	Female	0.02	0.02	0.00	0.00	5,853	38,426.25	24.46	38.58	1.32
4+Gender+M/H/T	✓	✓	✓	✓	T	Female	1.06	1.08	0.16	0.17	280,923	50,632.03	32.24	50.83	1.74
4+Gender+M/H/T	✓	✓	✓	✓	T	Male	1.41	2.50	0.22	0.38	374,563	50,856.98	32.38	51.05	1.74
4+Gender	✓	✓	✓	✓		Male	1.61	-	0.25	-	427,236	64,816.18	41.27	65.07	2.22
4+Gender+M/H/T	✓	✓	✓	✓	M	Male	0.07	2.56	0.01	0.39	17,558	71,887.86	45.77	72.17	2.46
4+Gender	✓	✓	✓	✓		Female	1.46	-	0.22	-	386,268	96,617.98	61.51	96.99	3.31
4+Gender+M/H/T	✓	✓	✓	✓	H	Male	0.13	2.69	0.02	0.41	35,115	210,178.50	133.81	211.00	7.21
4+Gender+M/H/T	✓	✓	✓	✓	H	Female	0.38	3.07	0.06	0.47	99,493	229,883.70	146.36	230.78	7.88

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,527 obs. in the urban areas and 25,149 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on income. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

Table 3.75: Most Opportunity Deprived Types (Accumulation of 5 Negative Circumstances) And Their Profiles For Per Capita Income, Urban Rural Area Level, **Rural Region**

											Rural Area				
Number of accumulated negative circumstances	Domi-nated Caste	Religious Minority	Father's with no education	Mother's with no education	Birth Region (M/ H/ T)	Gender	**Population share of sample** (total obs.)*	**Cumulative rural population share** (%)	**Estimated share of overall population** **	**Cumulative total population share** (%)	**Estimated population** ***	**Mean advantage**	**Ratio of rural area mean** (%)	**Ratio of national overall mean** (%)	**USD per day** ****
4+Gender+M/H/T	✓	✓	✓	✓	T	Male	0.99	0.99	0.84	0.84	263,375	49,739.60	55.72	49.93	1.71
4+Gender+M/H/T	✓	✓	✓	✓	H	Female	1.19	2.18	1.01	1.85	314,997	51,745.59	57.96	51.95	1.77
4+Gender+M/H/T	✓	✓	✓	✓	T	Female	0.94	3.12	0.80	2.65	248,626	52,345.66	58.64	52.55	1.79
4+Gender	✓	✓	✓	✓		Male	2.47	-	2.09	-	654,224	52,417.64	58.72	52.62	1.80
4+Gender+M/H/T	✓	✓	✓	✓	H	Male	1.24	4.36	1.05	3.69	327,639	53,518.67	59.95	53.73	1.84
4+Gender	✓	✓	✓	✓		Female	2.46	-	2.09	-	652,117	55,436.88	62.10	55.65	1.90
4+Gender+M/H/T	✓	✓	✓	✓	M	Male	0.24	4.60	0.20	3.90	63,210	57,869.14	64.82	58.09	1.98
4+Gender+M/H/T	✓	✓	✓	✓	M	Female	0.33	4.93	0.28	4.18	88,494	77,260.94	86.55	77.56	2.65

Note: a ✓ indicates that the most discriminated against category out of the categorical variables within that categorical variable is chosen. See regression and IOp analysis for more detail. *4,527 obs. in the urban areas and 25,149 observations in the rural areas. **Total of 29,588 obs. with 1) information on all circumstances and 2) information on income. ***Nepal's overall population is estimated at 26,494,504 people in 2011 according to the Central Bureau of Statistics (2012) ****av. exchange rate as of 13.11.2011 (1 USD = 79.90 NRs.)

3.6.2.4. Robustness Checks: Economies of Scale, Bootstrapping, Multicollinearity, OLS Regression Validity Checks, etc.

Great care is taken to ensure the most coherent, clean and robust inequality of outcome and IOp analyses. To do so, at every stage of the analytical process, a series of measures was taken to check the data for consistency, the methodology for coherence and the statistical tools and results for robustness.

Refer to the national level robustness checks section (3.4.2.3) for the general robustness checks during the various stages for the analyses. These were also carried out at the urban rural area level. The same categories are listed below. Only complementary and urban rural area level specific remarks are given here. The remaining are similar to the national level analyses.

1. **Dataset**: e.g. data cleaning for outliers, missing values (especially Section 3.3.2)

2. **Dependant variable aggregations**: e.g. data cleaning, outliers, literature guidance comparison, unit conversion checks (Section 2.4 for consumption and Section 2.5 for income)

3. **Inequality in outcome indices:** e.g. various different indices computations with various dependant variables, bootstrapped standard errors (Sections 2.2 and 2.6)

4. **Type distribution and circumstance categorical checks and variations:** e.g. low type distribution checks, circumstance categorical variable combination checks (mainly Sections 3.1.4 and 3.3.2.5)

Additionally to the national and Development Region level remarks, note that the sample population for the urban area is relatively low with only 46 out of 144 observed types, and of those, 20% with less than 4 observations per type. Despite this, the subsequent IOp, RF-OLS regression and partial IOR analyses are robust, with high levels of statistical significance, for most independent variables. Also, the R-squared values are much above what is generally found in the literature, particularly for the urban sample.

5. **Regressions:** e.g. various regression tests and various regressions (e.g. Link test, Ramsey test, Stepwise regressions, residual checks for normality and homoscedasticity, multicollinearity check, variance inflation factor, outlier and leverage checks, Dfbeta test), (RF-OLS regression sections for each level of analyses: Section 3.4.2.3 for the National level analyses, Section 3.5.2.4 for the Development Region level, Section 3.6.2.4 for the urban rural level, Section 3.7.2.4 for the grouping my population quarter level)

The presented regression results in the relevant section do not change drastically from one to another regression within each population sample. Most results are significant at the 1% level. R-squared is higher for the consumption than the income panel, and in fact for urban areas in the consumption panel it is the highest among all analyses in this study and higher than those of the literature. Compared to the literature, for the other urban rural area samples, the level of R-squared is seen as acceptable. In general, the consumption panel results are preferred over the income ones.

Despite the urban areas relatively low sample size for both economic outcome panels, the analyses' results appear valid.

A series of regression tests were carried out to test for their validity, homoscedasticity, etc. These also hold for this level's analyses. Refer to the national level robustness section for more details.

6. **IOp indices:** e.g. parametric and non-parametric, scale equivalence checks, axiom satisfaction checks, income and consumption expenditure analyses comparisons (IOp scalar indices and IOp equivalence scale sections for each level of analyses)

Table 3.76 presents the equivalence scale robustness checks results for this level's scalar IOp analyses. For both consumption expenditure and income and albeit some minor variations, the equivalence scale tests seem robust and consistent.

7. **Results**: e.g. comparison between different level analyses of this study, and to other empirical results in the literature (corresponding results interpretations sections for all level analyses, at both the descriptive and econometric level)

Where striking, the urban rural area level results are compared to the national and Development Region level ones. Attempts to interpret particular results are done, such as the parental education ones or the amplified importance for caste group belonging in urban areas.

Table 3.76: Equivalence Scale Robustness Checks on the Scalar Inequality of Opportunity Indices, Urban Rural Area Level

	Urban Region, all HH members*	**Equivalence Scale Application**		**Rural Regio,** all HH members**	**Equivalence Scale Application**	
		α=0.5	α=0.75		α=0.5	α=0.75
Per capita consumption expenditure (variable *pccons*)						
Total inequality (E0)	0.275	0.384	0.384	0.211	0.216	0.216
Non-parametric estimates						
IOL	0.121	0.176	0.176	0.046	0.036	0.036
IOR	0.442	0.458	0.458	0.219	0.166	0.166
Parametric estimates						
IOL	0.094	0.201	0.208	0.036	0.059	0.055
IOR	0.343	0.524	0.541	0.170	0.272	0.254
Per capita income (variable *pcincome*)						
Total inequality (E0)	0.464	0.604	0.604	0.550	0.594	0.594
Non-parametric estimates						
IOL	0.151	0.191	0.191	0.071	0.090	0.090
IOR	0.326	0.316	0.316	0.128	0.151	0.151
Parametric estimates						
IOL	0.085	0.200	0.224	0.041	0.088	0.094
IOR	0.184	0.332	0.371	0.074	0.147	0.159

Note: *For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.7. Inequality of Opportunity Analysis in Nepal: Population Grouping by Economic Welfare Quarters Level

This section (3.7) focuses on the empirical analyses of IOp by population income and consumption grouping by quarters. First, this section elaborates on the descriptive statistics.

Second, the empirical results are given and interpreted. To do so, initially the scalar indices of IOp, then the RF-OLS results and the partial IORs are given discussed. Following the circumstantial specific shares of unequal opportunities for the most opportunity deprived are highlighted and discussed. The descriptive and empirical findings of the population grouping level analyse are combined and fully discussed.

Finally, various robustness checks are carried out and discussed.

3.7.1. Descriptive Analysis: Income and Consumption Expenditure Grouping by Quarters

This section looks at some general noteworthy descriptive statistics that are different to the ones already elaborated and explained in the descriptive statistics section for the national, Development Region and the urban rural level analyses. It then discusses the type divisions and distributions for the income and consumption groupings by population quarters. Finally, the descriptive statistics for circumstance variables and their sub-categories are given with an initial attempt of interpretation.

The descriptive statistics are given for the bottom, second, third and top income and consumption quarter of the population, in other words, the earners and consumers. All seven circumstance variables also used at the Development Region and the national level analyses are utilised here. Note that all HH members who fall into the 30 to 49 age range and have complete information on all circumstances are included into the analyses. This is because the population grouping by quarter analyses is specific for this study, and one does not need to apply the HH head and spouse constraint for results comparability to the literature directly.

In general, and for the sample analysed, top earners have almost 70% of the national income share and top consumers about 56% (Table 3.78). For the **bottom quarters of the population,** earners and consumers hold less than 3% and 9% of the national total, respectively. The mean income level of the top earners is about 21 times that of the bottom quarter of the populations, while for consumers it is about 7 times that. These contrasts are stark, but not as extreme as in some more developed countries.

Of the **top earning and consuming quarters** of the population, more than 48% and 54%, respectively, live in the central region (Table 3.80). For both, earners and consumers, 39% of the third quarter also live in the central region and they hold an additional 20% or more percent of the national income.

At first glance, the bottom earner and consumer distribution across the five Development Regions appears relatively even. However, once the respective regional populations are accounted for, the far- and mid-west are almost exclusively populated by the **poorest of the poor**.

Table 3.84 presents the **type distributions** for income and consumption groups by population quarters. The higher the level of consumption and income, the higher the number of observed types. For instance, for the possible 288 type combinations only 92 types are observed for the bottom 25% consumers, while over 128 types are observed for the top 25% consumers. One explanation could be that combining the type distributions with the descriptive statistics from Table 3.79, is that the poorest 25% of the population are relatively homogenous in terms of their circumstances, and that predominantly with negative circumstances. In other words, they accumulate a set of exogenous

factors that predetermine their level of consumption and income. More on this on the econometric results interpretations of the next chapter.

The **sex ratio improves** and gets close to the national average the higher up one is in the income earning group. This can be a) because the richer households are more educated (positive correlation between mother's and father's education) and thus there being higher gender equality, and b) because particularly the poorer households depend on emigration and remittances in order generate enough economic revenue.

For consumption, apart from the top quarter (close to the national average of 89), the lower the consumption expenditure the higher the sex ratio reaching over 94 for the bottom quarter.

The descriptive statistics show the expected evolution as one moves from the bottom to the top earners and consumers. For example, a disproportionately larger share of individuals belong to the upper castes, have more educated parents, are Hindus, live in urban areas and in the hill region. Also, as mentioned in the previous descriptive statistics sections of the other analyses levels, richer families tend to live in the urban areas and tend to host family members for longer (i.e. sons and daughters stay beyond 30 years of age) and tend to host other family members to the family nucleus (e.g. children-in -law and siblings-in-law).

Interestingly, the proportion of **Buddhists** also increases as the level or earnings and consumption increases. More on this in the econometric results interpretation, but this is largely due to a) Buddhists conserving a relative elite status in Nepalese society, and b) the disconnection (or non-association) of Buddhists to the caste system, which is mainly linked to Hinduism, which in turn provides them with greater economic opportunities.

Table 3.77: Nominal Per Capita Consumption and Income by Population Quarters (no data constraints*)

	Per capita consumption expenditure (*pccons*)				**Per capita income** (*pcincome*)			
Quarters	Mean per capita consumption expenditure (NRs.)	Share of national consumption expenditure (%)	Cumulative share of consumption (%)	Obs.	Mean per capita income (NRs.)	Share of income (%)	Cumulative share of income (%)	Obs.
Poorest (First)	16,798.96	11.46	11.46	9,129	13,573.06	3.65	3.65	7,854
Second	28,229.10	16.72	28.18	7,923	42,045.10	10.99	14.64	7,631
Third	45,376.60	22.94	51.12	6,765	87,086.79	22.15	36.79	7,424
Top (Fourth)	111,937.60	48.88	100.00	5,843	269,603.00	63.21	100.00	6,843
Nepal mean	**45,112.63**	**100.00**		**29,660**	**98,106.84**	**100.00**		**29,752**

Note: * for all observations with complete information on all circumstance variables

Table 3.78: Nominal Per Capita Consumption and Income by Population Quarters (with data constraints*)

	Per capita consumption expenditure (*pccons*)				**Per capita income** (*pcincome*)			
Quarters	Mean per capita consumption expenditure (NRs.)	Share of national consumption expenditure (%)	Cumulative share of consumption (%)	Obs.	Mean per capita income (NRs.)	Share of income (%)	Cumulative share of income (%)	Obs.
Poorest (First)	**16,880.25**	**8.80**	8.80	1,859	**13,060.43**	**2.65**	2.65	1,636
Second	28,412.67	13.43	22.23	1,685	42,085.40	8.50	11.14	1,630
Third	45,625.82	22.05	44.29	1,723	88,129.71	19.51	30.66	1,788
Top (Fourth)	**110,938.40**	**55.71**	100.00	1,790	**275,417.60**	**69.34**	100.00	2,033
Nepal mean	**50,509.95**	**100.00**		**7,057**	**113,936.20**	**100.00**		**7,087**

Note: * for all observations 1) aged 30 - 49 years old, and 2) with complete information on all circumstance variables

Table 3.79: Definition and Description of Circumstance Variables, Income and Consumption Grouping by Quarter Level Analysis

	Bottom Quarter Consumers, Descriptive stats (%), all HH members*	**Second Quarter Consumers,** Descriptive stats (%), all HH members**	**Third Quarter Consumers,** Descriptive stats (%), all HH members***	**Top Quarter Consumers,** Descriptive stats (%), all HH members****	**Bottom Quarter Earners,** Descriptive stats (%), all HH members°	**Second Quarter Earners,** Descriptive stats (%), all HH members°°	**Third Quarter Earners,** Descriptive stats (%), all HH members°°°	**Top Quarter Earners,** Descriptive stats (%), all HH members°°°°
Gender								
Dominant (Male)	48.52	43.98	43.35	45.2	43.89	44.05	46.81	46.43
Dominated (Female)	51.48	56.02	56.65	54.8	56.11	55.95	53.19	53.57
Ethnicity								
Dominant caste/ethnic majority	21.25	26.71	36.33	**43.41**	27.81	25.46	29.08	42.84
Dominated caste/ethnic minority	**78.75**	73.29	63.67	56.59	72.19	74.54	70.92	57.16
Religion								
Hindu	83.49	83.38	82.76	85.81	83.13	82.82	82.72	86.18
Buddhist	7.1	9.08	**11.03**	**10.89**	7.52	8.65	11.02	10.58
Other	**9.41**	7.54	6.21	3.3	9.35	8.53	6.26	3.25
Father's education								
No schooling, illiterate, primary incomplete	83.81	74.12	63.96	**38.49**	78.3	71.35	67.9	47.07
Primary completed and above	**16.19**	25.88	36.04	61.51	21.7	28.65	32.1	52.93
Mother's Education								
No schooling, illiterate, primary incomplete	95.64	93.77	86.88	**67.99**	94.62	90.25	88.87	73.19
Primary completed and above	**4.36**	6.23	13.12	32.01	5.38	9.75	11.13	26.81
Birth region 1 (U/R)								
Urban	**3.93**	6.11	11.43	**28.72**	4.28	6.93	12.36	23.91
Rural	96.07	93.89	88.57	**71.28**	95.72	93.07	87.64	76.09
Birth region 2 (M/H/T)								
Mountain	8.61	10.5	9.63	8.99	11.06	10.49	8.05	8.36
Hill	51.05	50.03	61.29	**72.12**	54.77	50	60.51	67.19
Terai	**40.34**	39.47	29.08	18.88	34.17	39.51	31.43	24.45

Note 1: Mother's and father's education was divided into the two categories according to EducateNepal (2010) definitions and explanations. Categorical ethnical variables are divided as suggested by Bhattachan (2003 p.17) and Bennett et al (2008). Religion is divided as by Nepal Consensus Data 2011. Birth region 1 2 and gender are divided as suggested by NLSS III Report (CBS, 2011).
Note 2: Parental education was filled using the educational statistics for HHH/spouse, brothers/sisters and sons/daughters.
*For all 1,859 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 1,685 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
***For all 1,723 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
****For all 1,790 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°For all 1,1636 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°°For all 1,630 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°°°For all 1,788 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
°°°°For all 2,033 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.80: Birth Place Descriptive Statistics, Income and Consumption Grouping by Quarter Level (Variable *pccons* and *pcincome*)

Birth place Development Region	**Bottom Quarter Consumers,** Descriptive stats (%), all HH members*	**Second Quarter Consumers,** Descriptive stats (%), all HH members**	**Third Quarter Consumers,** Descriptive stats (%), all HH members***	**Top Quarter Consumers,** Descriptive stats (%), all HH members****	**Bottom Quarter Earners,** Descriptive stats (%), all HH members°	**Second Quarter Earners,** Descriptive stats (%), all HH members°°	**Third Quarter Earners,** Descriptive stats (%), all HH members°°°	**Top Quarter Earners,** Descriptive stats (%), all HH members°°°°
Eastern	23.24	26.94	26.93	17.21	24.14	26.32	23.88	20.61
Central	**18.56**	32.17	**39.18**	**54.47**	21.27	33.44	38.26	**47.71**
Western	17.48	17.98	21.42	21.45	17.18	19.82	19.63	21.15
Mid-western	23.56	14.18	7.54	4.41	21.58	12.09	12.14	**6.05**
Far-western	17.16	8.72	**4.93**	**2.46**	15.83	8.34	6.1	**4.48**
Total	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

Note: Author's computations using the NLSS III dataset for *pccons* and *pcincome*. Refer to Table 3.79 for footnote details.

Table 3.81: Relationship to Household Head, Consumption Grouping by Quarter Level Analyses (Variables *pccons*)

	Bottom Quarter Consumers Descriptive stats (%) all HH members*			**Second Quarter Consumers** Descriptive stats (%) all HH members**			**Third Quarter Consumers** Descriptive stats (%) all HH members***			**Top Quarter Consumers** Descriptive stats (%) all HH members****		
Relationship to HH head	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**
Household head	907	48.79	48.79	851	50.50	50.50	852	49.45	49.45	802	44.80	44.80
Husband/Wife	647	34.80	83.59	585	34.72	85.22	614	35.64	85.08	638	35.64	80.45
Son/Daughter	155	**8.34**	91.93	106	6.29	91.51	124	7.20	92.28	175	**9.78**	90.22
Brother/Sister	18	0.97	92.90	30	1.78	93.29	28	1.63	93.91	47	**2.63**	92.85
Father/Mother	11	0.59	93.49	3	0.18	93.47	13	0.75	94.66	8	0.45	93.3
Grandchild				1	0.06	93.53	1	0.06	94.72			
Nephew/Niece	2	0.11	93.60				1	0.06	94.78	1	0.06	93.35
Son/Daughter-In-Law	91	4.90	98.49	85	5.04	98.58	76	4.41	99.19	92	5.14	98.49
Brother/Sister-In-Law	24	1.29	99.78	12	0.71	99.29	7	0.41	99.59	20	**1.12**	99.61
Father/Mother-In-Law	1	0.05	99.84	1	0.06	99.35						
Other Family Relative	3	0.16	100	9	0.53	99.88	5	0.29	99.88	3	0.17	99.78
Servant/Servants Relatives				1	0.06	99.94	1	0.06	99.94	1	**0.06**	99.83
Other Non-Related				1	0.06	100	1	0.06	100	3	0.17	100
Total	**1,859**	**100**		**1,685**	**100**		**1,723**	**100**		**1,790**	**100**	

*For all 1'859 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 1,685 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ***For all 1,723 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ****For all 1,790 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.82: Relationship to Household Head, Income Grouping by Quarter Level Analyses (Variables *pcincome*)

	Bottom Quarter Earners Descriptive stats (%) all HH members*			**Second Quarter Earners** Descriptive stats (%) all HH members**			**Third Quarter Earners** Descriptive stats (%) all HH members***			**Top Quarter Earners** Descriptive stats (%) all HH members****		
Relationship to HH head	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**	**Freq.**	**Percent**	**Cum.**
Household head	816	49.88	49.88	854	52.39	52.39	868	48.55	48.55	884	43.48	43.48
Husband/Wife	588	35.94	85.82	506	31.04	83.44	642	35.91	84.45	757	37.24	80.72
Son/Daughter	100	**6.11**	91.93	116	7.12	90.55	149	8.33	92.79	202	**9.94**	90.65
Brother/Sister	26	**1.59**	93.52	21	1.29	91.84	37	**2.07**	94.85	41	**2.02**	92.67
Father/Mother	6	0.37	93.89	13	0.8	92.64	4	0.22	95.08	12	0.59	93.26
Grandchild				1	0.06	92.7				1	0.05	93.31
Nephew/Niece				3	0.18	92.88				1	0.05	93.36
Son/Daughter-In-Law	83	5.07	98.96	91	5.58	98.47	68	3.80	98.88	103	5.07	98.43
Brother/Sister-In-Law	13	0.79	99.76	18	1.1	99.57	11	0.62	99.5	22	**1.08**	99.51
Father/Mother-In-Law	1	0.06	99.82				1	0.06	99.55			
Other Family Relative	1	0.06	99.88	7	0.43	100	7	0.39	99.94	5	0.25	99.75
Servant/Servants Relatives							1	0.06	100	2	**0.10**	99.85
Other Non-Related	2	0.12	100							3	0.15	100
Total	**1,636**	**100**		**1,630**	**100**		**1,788**	**100**		**2,033**	**100**	

*For all 1,1636 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 1,630 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ***For all 1,788 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ****For all 2,033 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.83: Sex Ratio, Income and Consumption Grouping by Quarter Level Analyses

Sex ratio							
Bottom Quarter Consumers, Descriptive stats (%), all HH members*	**Second Quarter Consumers,** Descriptive stats (%), all HH members**	**Third Quarter Consumers,** Descriptive stats (%), all HH members***	**Top Quarter Consumers,** Descriptive stats (%), all HH members****	**Bottom Quarter Earners,** Descriptive stats (%), all HH members°	**Second Quarter Earners,** Descriptive stats (%), all HH members°°	**Third Quarter Earners,** Descriptive stats (%), all HH members°°°	**Top Quarter Earners,** Descriptive stats (%), all HH members°°°°
94.25	78.51	76.52	82.48	**78.22**	78.73	88.01	86.67

Note: Author's computations using the NLSS III dataset. Refer to Table 3.79 for footnote details.

Table 3.84: Type Division Description, Income and Consumption Grouping by Quarter Level Analyses

	Per capita consumption expenditure *(pccons)*				**Per capita income** *(pcincome)*			
	Bottom Quarter Consumers, Descriptive stats (%), all HH members*	**Second Quarter Consumers,** Descriptive stats (%), all HH members**	**Third Quarter Consumers,** Descriptive stats (%), all HH members***	**Top Quarter Consumers,** Descriptive stats (%), all HH members****	**Bottom Quarter Earners,** Descriptive stats (%), all HH members°	**Second Quarter Earners,** Descriptive stats (%), all HH members°°	**Third Quarter Earners,** Descriptive stats (%), all HH members°°°	**Top Quarter Earners,** Descriptive stats (%), all HH members°°°°
Maximum number of types	288	288	288	288	288	288	288	288
Number of types observed	**92**	101	133	**128**	**96**	112	129	**137**
Mean number of observations per type	20.20652	16.68	12.95	13.98	17.04	14.55	13.86	14.84
Min. number of observations per type	1	1	1	1	1	1	1	1
Max. number of observations per type	278	193	157	105	193	190	168	115
Proportion of types with fewer than 5 observations but more than 0.	17.36 (50)	17.36 (50)	23.61 (68)	19.79 (57)	17.00 (49)	19.79 (57)	23.61 (68)	19.44 (56)

Note: Author's computations using the NLSS III dataset for *pccons* and *pcincome*. Refer to Table 3.79 for footnote details.

3.7.2. Econometric Analysis for Inequality of Opportunity, Results and Results' Interpretation: Income and Consumption Expenditure Grouping by Quarters

This econometric analysis section consists first of the scalar indices of IOp. Second, it lays out and interprets the results of the RF-OLS regressions of the circumstance variables on economic outcome. It also looks closely at the circumstantial specific shares of unequal opportunities, the partial IORs. Third, the most opportunity deprived types are ranked by their average mean economic outcome and their profiles are interpreted. Last, the section highlights some of the robustness checks that were carried out throughout the analysis.

3.7.2.1. Parametric and Non-parametric Scalar Indices of Inequality of Opportunity

Table 3.85 presents the results for the scalar IOp indices for income and consumption groupings by population quarters. As expected, the total inequality in outcome (MLD or E_0) are higher for income than for consumption. However, since population the samples are grouped by economic outcome quarters, the level of total economic outcome inequality are relatively low, since the within quarter variances are relatively small and as expected the highest for the top earning and consuming quarters (refer to Table 3.78 for more details on the descriptive statistics of the population quarters).

For **consumption expenditure,** the higher the group mean consumption the higher the level of overall consumption inequality (E_0) and the higher the level of IOR (7% to 18% in the top quarter compared to 3% to 9% in the bottom quarter). In other words, people within the bottom 25% consumption quarter are all more or less equally poor (E_0 of only 0.03 compared to 0.1 for the top 25% consumers) and they all suffer from more or less the same degree of opportunity deprivation (3% to 9% of total inequality). In other words, the level of within group opportunity deprivation for the poorest 25% consumers is relatively low compared to that of the top 25% consumers. Top quarter consumers face higher consumption inequality (E_0 of 0.1) and the within group level of opportunity deprivation is much higher (7% to 18% of total consumption inequality). It is interesting to make assumptions when combining these results with those stated by the economic development facts in the country specific background section (3.3.1). In combination they suggest that for the limited economic (i.e. employment) opportunities available in Nepal, the most privileged in society are in fierce competition and they need to maximise the mobilisation of social capital. According to Ezemenari and Joshi (2019, p. 47) "Social status determines the types of personal networks individuals can access […] personal networks are an essential element for youth and adults who seek employment." This does not mean, however, that there is no between income and consumption type discrimination. As a matter of fact, the other analytical chapters emphasise the stark difference in opportunity deprivation levels when the grouping the population by varying social and ethnic characteristics. The goal of this section is to add depth to the IOp analysis and complement the previous results.

For **income**, total income inequality rises from the second to the top quarter of the population, as does the share of inequality of opportunity (between 3% to 23% of total income inequality). This is consistent with the consumption IOp indices. However, for the bottom income quarter of the population, the level of opportunity deprivation out of overall income inequality (relatively high, 0.11) ranges from 2% to 54%. There are some observations in the sample that have a large negative revenue, which in return bias the estimation results. Since great care was taken in the aggregation process for the income variable (refer to section 2.5), these negative values were deemed to be acceptable and plausible. This was also confirmed during various meetings with individuals at various public and

private institutions during the author's field trip to Nepal (October 2018). It was concluded that a large share of emigrants need to indebt themselves in order to be able to emigrate and remaining HH members live in uncertainty if they will receive remittances. These people seem to be trapped in a downward borrowing spiral.

This assumption is backed by the equivalence scale measures where with α below 1 the extreme negative values lose in significance and opportunity deprivation ranges between 22% to 35%. More on this in Section 3.7.2.4.

Table 3.95 reports the results of the paired t-test to test for **statistical difference** between all levels of the IOp analyses, i.e. national, Development Region, urban rural area, and by income/consumption population quarter level. The correlation matrix is available upon request. The coloured boxes within the table show the within level results statistical difference comparison. At the **within consumption and income by population quarter level** most of the computed IOp indices are not statistically different between the different outcome quarter groups. The only statistical difference is between the second, third and top consumer quarters.

Table 3.85: Scalar Indices of Inequality of Opportunity, Income and Consumption Grouping by Quarter Level Analyses

	Bottom Quarter Earners/Consumers, Descriptive stats (%), all HH members*	**Second Quarter Earners/Consumers,** Descriptive stats (%), all HH members**	**Third Quarter Earners/Consumers,** Descriptive stats (%), all HH members***	**Top Quarter Earners/Consumers,** Descriptive stats (%), all HH members****
Per capita consumption expenditure (variable *pccons*)				
Total inequality (E0)	0.028	0.008	0.012	0.096
Non-parametric estimates				
IOL	0.003	0.001	0.002	0.017
IOR	**0.089**	0.124	0.173	**0.180**
Parametric estimates				
IOL	0.001	0.000	0.000	0.007
IOR	**0.031**	0.032	0.040	**0.072**
Per capita income (variable *pcincome*)				
Total inequality (E0)	0.106	0.025	0.023	0.148
Non-parametric estimates				
IOL	0.060	0.002	0.004	0.033
IOR	**0.535**	0.077	0.174	**0.226**
Parametric estimates				
IOL	0.000	0.000	0.001	0.005
IOR	**0.017**	0.007	0.037	**0.035**

Note: *For all 1,859 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
**For all 1,685 obs. for per capita consumption expenditure and 1,1630 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
***For all 1,723 obs. for per capita consumption expenditure and 1,788 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
****For all 1,790 obs. for per capita consumption expenditure and 2,033 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.7.2.2. Reduced-Form OLS Regressions and Circumstance Specific Shares of Opportunities (Partial IORs)

The RF-OLS regression results for the economic outcome by population grouping by quarters on per capita consumption expenditure are reported in Table 3.87 and on per capita income in Table

3.88. For comparison purpose to the national level, Development Region level and urban rural level analyses, the same regression and categorical variable combinations are run here. The correlation matrices are available upon request.

Table 3.86 shows the partial IOR estimates, that is the circumstance specific shares of overall opportunities compared to overall inequality in outcome. The results of this level's analyses have to be regarded with even more caution than the other levels. This is because numerous circumstances are statistically insignificant as indicated by the RF-OLS regressions and because overall IOR (parametric) are very low for all population groupings across both outcome panels (only between 1% to 7%).

For both consumption and income, a lot of the regression results are statistically insignificant. R-squared is also relatively low with values below 0.05. This is not surprising since the regressions test for the impact of the categorical circumstance variables on economic outcome, which in turn is pre-selected and have relatively low levels of variance within each quarter of economic outcome (refer to the descriptive section for more details).

It is interesting to note that in the consumption panel for both the bottom and top quarter of the population, only father's and not mother's education level is statistically significant. For income, mother's education is more significant and more important than father's education in the top quarter, whilst it is the reverse for the bottom quarter. These results may suggest that parental education is still a key driving force for causing the limited variance within the economic welfare groupings. These results are largely confirmed by the partial IOR estimates. For the consumption panel, they suggest that father's education and for the income that mother's education accounts for a larger share of opportunities. This is somewhat consistent with the literature on the measurement of economic welfare and prevalence of mother's education, especially given Nepal's developing country context.

Furthermore, for the top consumers being born in a rural area and belonging to a minority caste is statistically significant at the 1% level and negatively impacts on consumption. This is as expected. For the other consumption population quarters, however, rural is also significant but only at the 5% level and caste belonging only at the 5% or 10% level depending on the regression referred to.

Table 3.86: Circumstantial Specific Opportunity Share (Partial IORs), Consumption and Income Quarter Level

	Bottom Quarter Earners & Consumers (all HH members)'	**Second Quarter Earners & Consumers** (all HH members)"	**Third Quarter Earners & Consumers** (all HH members)"'	**Top Quarter Earners & Consumers** (all HH members)""
Per capita consumption expenditure (variable *pccons*)				
Mean Logarithmic Deviation	0.028	0.008	0.012	0.096
IOp Index Ratio (parametric)	0.031	0.032	0.040	0.072
Circumstances				
Ethnic Minority / Dominated Caste	-	0.008	0.017	0.026
Buddhist	-	-	-	0.012
Other Religion	0.050	-	0.001	-
Father's Education: Primary Complete or Above	0.035	0.015	-	0.037
Mother's Education: Primary Complete or Above	-	0.006	0.020	0.003
Rural Birth Region	0.035	0.007	0.014	0.033
Hill Birth Region	0.009	0.011	0.012	0.015
Terai Birth Region	0.034	-	-	0.033
Female	0.034	-	-	-
Per capita income (variable *pcincome*)				
Mean Logarithmic Deviation	0.106	0.025	0.023	0.148
IOp Index Ratio (parametric)	0.017	0.007	0.037	0.035
Circumstances				
Ethnic Minority / Dominated Caste	0.032	-	0.012	0.003
Buddhist	0.055	-	0.026	-
Other Religion	-	0.010	-	0.000
Father's Education: Primary Complete or Above	0.067	-	0.008	0.008
Mother's Education: Primary Complete or Above	-	-	0.010	0.014
Rural Birth Region	0.064	-	0.005	-
Hill Birth Region	-	-	0.019	-
Terai Birth Region	0.065	-	0.011	-
Female	-	-	-	0.010

Note: only circumstances that were statistically significant at least at the 10 percent level were used to compute the partial IORs.
'For all 1,859 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
"For all 1,685 obs. for per capita consumption expenditure and 1,1630 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
"'For all 1,723 obs. for per capita consumption expenditure and 1,788 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.
""For all 1,790 obs. for per capita consumption expenditure and 2,033 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.87: Reduced-Form OLS Regressions of Circumstances on Per Capita Consumption Expenditure, Consumption Grouping by Quarter Region Level

Circumstances	Bottom Quarter Consumers ln(pccons)				Second Quarter Consumers ln(pccons)				Third Quarter Consumers ln(pccons)				Top Quarter Consumers ln(pccons)			
Ethnic Minority / Dominated Caste	-0.015	-0.014	0.022	0.023	-0.015*	-0.016*	-0.031***	-0.031***	-0.015	-0.015	-0.020**	-0.019**	**-0.084*** **	-0.084***	-0.094***	-0.093***
	(0.016)	(0.016)	(0.015)	(0.015)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.023)	(0.023)	(0.022)	(0.022)
Buddhist	-0.006	-0.006	-0.028	-0.028	-0.021	-0.02	-0.009	-0.009	0.02	0.02	0.023	0.023	0.046	0.045	0.060*	0.060*
	(0.027)	(0.026)	(0.026)	(0.026)	(0.014)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.015)	(0.015)	(0.031)	(0.031)	(0.031)	(0.031)
Other Religion	**-0.072** **	-0.072**	-0.086***	-0.086***	-0.01	-0.01	-0.008	-0.008	-0.028*	-0.028*	-0.027	-0.027	0.091	0.091	0.087	0.087
	(0.029)	(0.029)	(0.029)	(0.029)	(0.012)	(0.012)	(0.013)	(0.013)	(0.017)	(0.017)	(0.017)	(0.017)	(0.065)	(0.066)	(0.062)	(0.063)
Father's Education: Primary Complete or Above	**0.101*** **	0.099***	0.104***	0.102***	0.013	0.013*	0.012	0.012	0.011	0.011	0.013	0.013	**0.062*** **	0.062***	0.061***	0.061***
	(0.013)	(0.013)	(0.014)	(0.014)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.021)	(0.021)	(0.021)	(0.021)
Mother's Education: Primary Complete or Above	0.007	0.011	0.012	0.017	0.020*	0.019*	0.017	0.016	0.031**	0.033**	0.029**	0.031**	0.034	0.036*	0.031	0.033
	(0.024)	(0.024)	(0.024)	(0.024)	(0.012)	(0.011)	(0.012)	(0.011)	(0.013)	(0.013)	(0.013)	(0.013)	(0.022)	(0.022)	(0.022)	(0.022)
Rural Birth Region	**-0.049** **	-0.051**	-0.069***	-0.072***	-0.028*	-0.028*	-0.025	-0.025	-0.034**	-0.034**	-0.034**	-0.034**	**-0.200*** **	-0.200***	-0.205***	-0.205***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.015)	(0.015)	(0.016)	(0.016)	(0.014)	(0.014)	(0.015)	(0.015)	(0.024)	(0.024)	(0.023)	(0.023)
Hill Birth Region	0.063**	0.062**			0.021*	0.022*			0.029*	0.029*			-0.029	-0.03		
	(0.025)	(0.025)			(0.012)	(0.012)			(0.016)	(0.016)			(0.032)	(0.032)		
Terai Birth Region	**0.113*** **	0.113***			-0.012	-0.012			0.006	0.006			-0.074**	-0.073**		
	(0.026)	(0.026)			(0.013)	(0.013)			(0.017)	(0.017)			(0.037)	(0.037)		
Female	-0.019		-0.021*		0.005		0.006		-0.008		-0.007		-0.015		-0.014	
	(0.012)		(0.012)		(0.007)		(0.007)		(0.008)		(0.008)		(0.020)		(0.020)	
Constant	9.672***	9.664***	9.744***	9.735***	10.265***	10.267***	10.275***	10.279***	10.715***	10.711***	10.736***	10.731***	11.626***	11.617***	11.597***	11.589***
	(0.035)	(0.035)	(0.026)	(0.025)	(0.021)	(0.021)	(0.018)	(0.017)	(0.023)	(0.022)	(0.017)	(0.017)	(0.045)	(0.044)	(0.035)	(0.032)
Observations	1859	1859	1859	1859	1685	1685	1685	1685	1723	1723	1723	1723	1790	1790	1790	1790
R-squared	0.057	0.056	0.041	0.039	0.031	0.031	0.019	0.018	0.026	0.026	0.02	0.02	0.074	0.073	0.071	0.071
R-squared (adjusted)	0.053	0.051	0.038	0.036	0.026	0.026	0.015	0.015	0.021	0.021	0.016	0.016	0.069	0.069	0.067	0.068
F stat	10.703	11.716	10.918	12.273	5.3	5.945	4.145	4.713	4.684	5.057	4.231	4.669	14.631	16.297	17.705	20.501

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,859 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,685 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,723 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita consumption aggregation methodology proposed by the author (variable *pccons*).
Data sample constraints: For all 1,790 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Table 3.88: Reduced-Form OLS Regressions of Circumstances on Per Capita Income, Income Grouping by Quarter Region Level

Circumstances	Bottom Quarter Earners ln(pcincome)				Second Quarter Earners ln(pcincome)				Third Quarter Earners ln(pcincome)				Top Quarter Earners ln(pcincome)			
Ethnic Minority / Dominated Caste	-0.096**	-0.096**	-0.043	-0.042	0.012	0.012	0.022	0.022	-0.01	-0.009	-0.028*	-0.028*	-0.059**	-0.059*	-0.061**	-0.061**
	(0.037)	(0.037)	(0.033)	(0.033)	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.030)	(0.030)	(0.029)	(0.029)
Buddhist	0.167***	0.167***	0.117**	0.117**	-0.009	-0.01	-0.016	-0.016	0.046**	0.045**	0.068***	0.068***	-0.018	-0.015	-0.016	-0.014
	(0.052)	(0.052)	(0.047)	(0.047)	(0.027)	(0.027)	(0.026)	(0.026)	(0.022)	(0.022)	(0.022)	(0.022)	(0.043)	(0.043)	(0.041)	(0.042)
Other Religion	-0.069	-0.07	-0.087	-0.087	-0.055***	-0.056***	-0.059***	-0.060***	0.009	0.009	0.001	0.001	-0.155***	-0.153***	-0.153***	-0.151***
	(0.070)	(0.070)	(0.073)	(0.073)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)	(0.045)	(0.045)	(0.044)	(0.044)
Father's Education: Primary Complete or Above	0.077*	0.077*	0.082**	0.082**	-0.001	0	0.001	0.001	0.025*	0.025*	0.025*	0.025*	0.062**	0.062**	0.062**	0.061**
	(0.040)	(0.041)	(0.041)	(0.041)	(0.016)	(0.016)	(0.016)	(0.016)	(0.014)	(0.014)	(0.014)	(0.014)	(0.028)	(0.028)	(0.028)	(0.028)
Mother's Education: Primary Complete or Above	0.037	0.038	0.045	0.047	-0.015	-0.016	-0.014	-0.015	0.051**	0.052***	0.046**	0.047**	0.097***	0.102***	0.096***	0.102***
	(0.062)	(0.062)	(0.063)	(0.063)	(0.025)	(0.025)	(0.025)	(0.025)	(0.020)	(0.020)	(0.020)	(0.020)	(0.030)	(0.030)	(0.030)	(0.030)
Rural Birth Region	-0.142***	-0.142***	-0.169***	-0.169***	0.017	0.017	0.016	0.016	-0.037**	-0.037**	-0.038**	-0.038**	0.001	0.004	0.002	0.005
	(0.048)	(0.048)	(0.049)	(0.049)	(0.027)	(0.027)	(0.027)	(0.027)	(0.018)	(0.018)	(0.019)	(0.019)	(0.031)	(0.031)	(0.030)	(0.030)
Hill Birth Region	0.028	0.028			0.007	0.007			-0.052**	-0.053**			-0.01	-0.012		
	(0.058)	(0.058)			(0.022)	(0.022)			(0.025)	(0.025)			(0.039)	(0.039)		
Terai Birth Region	0.136**	0.136**			0.022	0.023			-0.094***	-0.094***			-0.014	-0.015		
	(0.064)	(0.064)			(0.023)	(0.023)			(0.026)	(0.026)			(0.043)	(0.043)		
Female	-0.003		-0.007		0.009		0.009		-0.006		-0.004		-0.043*		-0.043*	
	(0.031)		(0.031)		(0.013)		(0.013)		(0.012)		(0.012)		(0.026)		(0.026)	
Constant	9.705***	9.704***	9.764***	9.761***	10.584***	10.588***	10.591***	10.596***	11.435***	11.432***	11.382***	11.380***	12.368***	12.343***	12.358***	12.332***
	(0.074)	(0.071)	(0.056)	(0.054)	(0.037)	(0.036)	(0.031)	(0.030)	(0.033)	(0.033)	(0.024)	(0.023)	(0.054)	(0.051)	(0.047)	(0.042)
Observations	1603	1603	1603	1603	1630	1630	1630	1630	1788	1788	1788	1788	2033	2033	2033	2033
R-squared	0.022	0.022	0.014	0.014	0.008	0.008	0.007	0.007	0.035	0.035	0.023	0.022	0.029	0.027	0.029	0.027
R-squared (adjusted)	0.016	0.017	0.01	0.01	0.002	0.003	0.003	0.003	0.031	0.031	0.019	0.019	0.025	0.023	0.026	0.024
F stat	3.371	3.792	3.697	4.306	1.142	1.247	1.346	1.515	5.346	5.998	4.629	5.387	6.693	7.27	8.568	9.65

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 71,603 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.
Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 1,630 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses. Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 1,788 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses. Statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1

Note: using the per capita income aggregation methodology proposed by the author (variable *pcincome*).
Data sample constraints: For all 2,033 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.
Robust standard errors in parentheses.

3.7.2.3. Most Opportunity Deprived Types and Their Profiles

No computation of the most opportunity deprived types is done for this level analyses. This is because population samples here are stratified by income and consumption expenditure quarters. The necessary relevant and same information can be obtained from the national level analyses.

3.7.2.4. Robustness Checks: Economies of Scale, Bootstrapping, Multicollinearity, OLS Regression Validity Checks, etc.

Great care is taken to ensure the most coherent, clean and robust inequality of outcome and IOp analyses. To do so, at every stage of the analytical process, a series of measures was taken to check the data for consistency, the methodology for coherence and the statistical tools and results for robustness.

Refer to the national level robustness checks section (3.4.2.3) for the general robustness checks during the various stages for the analyses. These were also carried out at the economic outcome grouping by quarters level. The same categories are listed below. Only complementary remarks concerning this level's analyses are given here. The remaining main elements are similar to the national level analyses.

1. **Dataset**: e.g. data cleaning for outliers, missing values (especially Section 3.3.2)

2. **Dependant variable aggregations**: e.g. data cleaning, outliers, literature guidance comparison, unit conversion checks (Section 2.4 for consumption and Section 2.5 for income)

3. **Inequality in outcome indices:** e.g. various different indices computations with various dependant variables, bootstrapped standard errors (Sections 2.2 and 2.6)

4. **Type distribution and circumstance categorical checks and variations:** e.g. low type distribution checks, circumstance categorical variable combination checks (mainly Sections 3.1.4 and 3.3.2.5)

Additionally to the national level remarks, note that the sample population for all three groupings are almost equal. Yet, the lower the mean level of consumption or income by population quarter the fewer types are observed.

5. **Regressions:** e.g. various regression tests and various regressions (e.g. Link test, Ramsey test, Stepwise regressions, residual checks for normality and homoscedasticity, multicollinearity check, variance inflation factor, outlier and leverage checks, Dfbeta test), (RF-OLS regression sections for each level of analyses: Section 3.4.2.3 for the National level analyses, Section 3.5.2.4 for the Development Region level, Section 3.6.2.4 for the urban rural level, Section 3.7.2.4 for the grouping my population quarter level)

For both economic outcome panels, a lot of the regression results are statistically insignificant. R-squared is also relatively low with values below 0.05. This is not surprising since the regressions test for the impact of the categorical circumstance variables on economic outcome, which in turn is

pre-selected and how relatively low levels of variance within each quarter of economic outcome (refer to the descriptive section for more details). Nevertheless, the results revealed a few interesting results.

6. **IOp indices:** e.g. parametric and non-parametric, scale equivalence checks, axiom satisfaction checks, income and consumption expenditure analyses comparisons (IOp scalar indices and IOp equivalence scale sections for each level of analyses)

Table 3.90 presents the equivalence scale robustness checks results for this level's scalar IOp analyses. For both economic outcome panels, the results for the non-parametric estimates are consistent, as expected (albeit with minor variations) and robust. The equivalence scale measures even "correct" the outlier effects of the negative income observations in the bottom earning quarter. As a result, the IOp estimates lose their absurdly high values and go back to a more likely level. It is suggested, however, to look at the bottom consumers rather than bottom earners IOp estimates.

An additional sensitivity check of IOp estimates to extreme values of outcome and variation is done. For this, the suspicious results from the income by population quarter analyses were chosen and in particular for the bottom 25% of earners. Table 3.89 compares the computed MLD, IOL and IOR indices before and after having deleted the observations for which a negative level of income was aggregated (see Chapter 2). Section 2.6 shows high negative values for minimum income observations for the bottom quintile as well as large standard deviation and Coefficient of Variation. Despite these seemingly large values, outliers were individually judged during the aggregation process. It is the accumulation of a series of plausible negative values that lead to such extreme aggregate values. Thus the overall aggregates are accepted to be that negative. Note that there is no information on saving levels of families to see whether they can afford this negative income.

Anyhow, the previously computed and extremely high level of IOR (non-parametric) almost disappears and the MLD estimate drastically shrink after deleting the negative income observations. This indicates that the estimators are sensitive to extreme values, particularly in the non-parametric approach which does not impose functional forms. However, this result highlights the importance of the complementarity between both the parametric and non-parametric approaches. In both approaches and no matter whether the negative income sources are included or excluded, the IOL and MLD estimates are relatively low indicating that opportunity deprivation within type (bottom earners) is minimal.

For both panels, the IOp indices particularly for the parametric estimates appear to be sensitive to the scale measures.

7. **Results**: e.g. comparison between different level analyses of this study, and to other empirical results in the literature (corresponding results interpretations sections for all level analyses, at both the descriptive and econometric level)

Where striking, the economic outcome by population grouping results are compared to those of the other level ones. Attempts to interpret particular results, such as the parental education ones, are done.

Table 3.89: IOp Estimates Before and After Eliminating Negative Income Observations

		Bottom Quarter Earners **BEFORE** deleting negative income observations*	**Bottom Quarter Earners** **AFTER** deleting negative income observations**
Per capita income (variable *pcincome*)			
Total inequality (E0)		0.112	0.026
Non-parametric estimates			
	IOL	0.060	0.002
	IOR	0.535	0.094
Parametric estimates			
	IOL	0.000	0.000
	IOR	0.001	0.009

Note: Descriptive stats (%), all HH members. *For all 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 1,594 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.90: Equivalence Scale Robustness Checks on the Scalar Inequality of Opportunity Indices, Income and Consumption Grouping by Quarter Level Analyses

	Bottom Quarter Earners/ Consumers, Descriptive stats (%), all HH members*	**Equivalence Scale Application**		**Second Quarter Earners/ Consumers,** Descriptive stats (%), all HH members**	**Equivalence Scale Application**		**Third Quarter Earners/ Consumers,** Descriptive stats (%), all HH members***	**Equivalence Scale Application**		**Top Quarter Earners/ Consumers,** Descriptive stats (%), all HH members****	**Equivalence Scale Application**	
		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75		α=0.5	α=0.75
Per capita consumption expenditure (variable *pccons*)												
Total inequality (E0)	0.028	0.081	0.081	0.008	0.095	0.095	0.012	0.083	0.083	0.096	0.227	0.227
Non-parametric estimates												
IOL	0.003	0.013	0.013	0.001	0.020	0.020	0.002	0.014	0.014	0.017	0.078	0.078
IOR	0.089	0.166	0.166	0.124	0.211	0.211	0.173	0.172	0.172	0.180	0.344	0.344
Parametric estimates												
IOL	0.001	0.047	0.055	0.000	0.070	0.083	0.000	0.056	0.068	0.007	0.124	0.138
IOR	0.031	**0.579**	**0.678**	0.032	**0.738**	**0.882**	0.040	**0.676**	**0.822**	0.072	**0.545**	**0.608**
Per capita income (variable *pcincome*)												
Total inequality (E0)	0.106	0.178	0.178	0.025	0.120	0.120	0.023	0.120	0.120	0.148	0.261	0.261
Non-parametric estimates												
IOL	0.060	**0.039**	0.039	0.002	0.025	0.025	0.004	0.021	0.021	0.033	0.055	0.055
IOR	0.535	**0.220**	0.220	0.077	0.207	0.207	0.174	0.179	0.179	0.226	0.212	0.212
Parametric estimates												
IOL	0.000	**0.052**	0.062	0.000	0.073	0.088	0.001	0.074	0.091	0.005	0.093	0.110
IOR	0.017	**0.291**	0.351	0.007	**0.605**	**0.734**	0.037	**0.622**	**0.758**	0.035	0.356	0.421

Note: *For all 1,859 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 1,685 obs. for per capita consumption expenditure and 1,1630 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ***For all 1,723 obs. for per capita consumption expenditure and 1,788 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ****For all 1,790 obs. for per capita consumption expenditure and 2,033 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

3.8. Inequality of Opportunity Results' Comparisons and How They Fit Into the General Existing Literature

This section combines the main econometric results of the IOp analyses across all four analysed levels. It further compares these to the results in the literature. The goal is to give a general overview of all the results obtained. Please note that there are some overlaps to the results interpretation sections already given at each level before, but an effort is made to minimise these. These sections go into much more detail and interpretation of their respective results. Thus when lacking here, further explanations can be obtained in the relevant sections. A condensed version of the key and counterintuitive results and the main literature contributions can be read in Table 4.1 in the next Chapter (4).

Table 3.93 summarises the main inequality in outcome and IOp indices and some other descriptive statistics for the national level analyses and the subsequent disaggregations. This section consistently follows the same structure as each of the four analysed levels.

1) Descriptive statistics and data constraints
2) IOp, other inequality indices and statistical difference tests
3) Regression analyses, correlation matrices and partial IORs
4) Opportunity deprivation share of circumstances and poverty
5) Robustness checks

1) Descriptive statistics and data constraints:

The mean **HH size** at the national level is 4.79 family members. Across all descriptive statistics at all levels of the analyses, it is evident that the practice of joint-families is still common in Nepal. There are variations according to geographical areas and division of population shares by mean income and consumption.

The national **sex ratio** is at 89.9 men to 100 women. This varies according to geographical region and economic outcome distribution. For instance, urban versus rural areas have 100 against 77 men per 100 women. It appears that the further away the Development Region from the capital, the lower the sex ratio. For instance, the far-western region has the lowest sex ratio of below 65, while the central region the highest with a sex ratio almost equal to the national average.

Type sample partitions varies, largely due to variations in sample size. The descriptive statistics suggest that the smaller the sample size the further away a region from the capital, and the more categorical variables, the lower the type distribution. See the type distribution section for a closer debate on the trade-off between the number of circumstances and type distribution.

Circumstance variables: The population **shares of the categorical circumstance variables** vary largely according to the geographical level of analysis. For instance, at the national level about 69% of the population belong to the dominated caste group. However, when the data is disaggregated by Development Region only 38% of the population in the far-west, yet 66% of people in the rural area belong to the dominated caste groups. Mother's education level shows even stronger distributional contrasts across the country. At the national level, over 24% of all mothers have at least completed primary education. In urban areas, this is over 32%, whilst in the far-west its only 9%. The share of mother's with basic education in the bottom 25% consumption group is only 5%.

Geographical mobility: Looking at the descriptive statistics, one can conclude that there is great domestic (and international) migration. For instance, the mid-western and far-western regions

have about half as many sons and daughters living with them than the other regions. There, also the sex ratios are particularly low (between 60 to 70). Combined with the caste statistics, it can be suggested that particularly the people of dominated caste group belonging tend to migrate to urban centres and particularly to the capital area in the central region. As low as 38% of the remaining population in the far-west belong to the dominated caste group (national average is 70%).

2) IOp, other inequality indices and statistical difference tests

There are **important variations in terms of mean per capita consumption expenditure and income** across the country. At the national level, mean per capita consumption expenditure is about 630USD/year. However, in the far-western region the mean level is at less than 380USD, whilst in the central region it is almost 800USD. In the urban areas, the mean per capita consumption is almost double that of the rural areas (1,100USD and 570USD, respectively. The poorest 25% of the population have a mean per capita consumption expenditure level of only 210 USD/year, whilst the top 25% consumers reach levels of over 6 times that (1400USD).

Inequality in outcome estimates: Overall, the inequality in income estimates for the Gini and the MLD estimates are higher than those for consumption expenditure. This is consistent with the literature.

IOp in income vs. consumption estimates: As expected, the levels of IOL for income are consistently higher to those of consumption. This is because the income measurement error and income variance are higher than those for consumption expenditure. In turn, this results in the IOR estimates for consumption to be higher than those for income (mostly). In other words, the residual inequality in the income distribution is larger (FG, 2011). This is also consistent with the literature and the results obtained by FG (2011).

Parametric vs. non-parametric estimates: Consistently, the non-parametric estimates of IOp are higher than those of the parametric estimates. This is as expected since for the parametric analyses, a variety of data distribution constraints are applied (distributions are smoothened and/or standardised), resulting in less between and within type variance. This results in lower overall IOL and IOR measures. This is also consistent with the literature and the results obtained by FG (2011).

At the **national level**, the range of opportunity deprivation out of total consumption inequality reaches from 26% to 32%. This is relatively moderate compared to other literature (e.g. FG, 2011) and relatively similar to developed countries (e.g. Checchi and Peragine, 2010). The suggestion here is that the relatively low level of IOp (national level) is due to Nepal's extremely low level of development (see Section 3.10.2.4: The Kuznets Curve Revisited. for a detailed explanation). However, once the national level results are disaggregated geographically by Development Region and urban versus rural areas, the results indicate a **large divergence in IOp estimates across the country**.

At the **Development Region level**, the central region faces the highest opportunity deprivation rates ranging from 37% to 44% of total inequality in consumption expenditure, while the mid-western region "only" faces 17% to 29% of opportunity depravation.

The disaggregation of the national level into **urban and rural** area also highlights a significant divergence in IOp across the country. Urban areas suffer from 34% to 44% and rural areas "only" from 17% to 21% of opportunity deprivation out of overall consumption inequality. Individuals tend to migrate to the central region and urban areas (especially the capital Kathmandu in the central region)

hoping for more opportunities. However, these are the areas with highest levels of opportunity deprivation due to unfavourable circumstances. Furthermore, from the descriptive and regression sections one can conclude that the more negative circumstances one accumulates the more one is discriminated against in the central and the urban regions.

Income and consumption quarters: Disaggregating IOp not only by geographical factors but also by population income and consumption quarters adds significant depth and layers of understanding to the study of IOp in Nepal. (Income estimates for the bottom quarter of the population can be disregarded due to outliers. Reference made to consumption expenditure only). The poorest 25% of the population in terms of consumption expenditure, among themselves, suffer from relatively little opportunity deprivation (3% to 9%), whilst the **top quarter of the population** faces opportunity deprivation of 7% to 18% out of total consumption inequality. While this result may seem contradictory to the rest of this study, they are not. Combining them to the descriptive economic development statistics from section 3.3.1, they suggest that for the **limited opportunities available in Nepal**, the most privileged in society are in fierce competition and need to maximise the mobilisation of social capital to seize them. Social networks do not only act as a means of socio-economic safety nets (i.e. insurance), but are also "an essential element for youth and adults who seek employment and the quality of these networks are often linked to levels of income and social status" (Ezemenari and Joshi, 2019, p. 6).

There is an imminent **link between the IOp indices calculated in this study and the level of accessibility to** services, employment and so on, in other words opportunity to access. Inevitably, a lack of alternative, such as in the remote rural regions, contributes to the apparent deprivate computed in this study. UNDP and the National Planning Commission (2014, Ch. 3) have dedicated an entire chapter of their Nepal Human Development Report to estimate regional differences in access to facilities (e.g. drinking water, education institutions, health care centres). Furthermore they compute a HH Well-Being Index in to be able to correlate access to the livelihoods of HHs, also depending on their caste and ethnic group belonging. They also use the NLSS dataset rounds to compute their results. Similarly to the results of this study, they find that HH well-being varies largely according to ethnic or caste group belonging, emphasizing the urgent need to close regional development gaps to enhance well-being. Stiglitz (2018) emphasises that greater inclusivity can lead to more robust growth, and that economic growth and equality act complementarily.

Ezemenari and Joshi (2019) highlight that inequality and its disparity across social, income and geographic dimensions significantly hinders individuals and HHs to perform to their fullest potential. The results of their adjusted Human Capital Indices to socio-economic and ethnic elements suggest that income and social disparities are strongly correlated with human capital outcomes, which vary also greatly geographically. These inequalities limit the accessibility to quality services for human capital development. They state that "improved human capital outcomes will require better quality of services and access to services (which often depend as well on the socioeconomic status of the beneficiaries)."

The **statistical differences** between the **IOL versus IOR, parametric versus non-parametric estimates** were tested. Table 3.91 shows all the results using the paired t-test. In all cases the null can be rejected and the means of the IOp estimates, no matter whether IOL, IOR, parametric or non-parametric, are different from each other.

Furthermore, the statistical differences between the IOL and IOR estimates (parametric and non-parametric, for income and consumption) **between all levels of the IOp analyses** were tested, i.e. national, Development Region, urban rural area, and by income/consumption population quarter level. Table 3.95 reports the results of the paired t-test. The correlation matrix is available upon request. The

coloured boxes within the table show the within level results' statistical difference comparison. Several remarks can be made

i. **Within Development Region level**: the computed IOp indices are statistically different between the eastern, central and western regions. There is no statistical difference between the IOp estimates with the far- and mid-western regions.

ii. **Within urban rural area level**: the computed IOp indices are statistically different between the urban and the rural areas.

iii. **Within consumption/income by population quarter level**: most of the computed IOp indices are not statistically different between the different outcome quarter groups. The only statistical difference is between the second, third and top consumer quarters.

iv. **Between national and the other three levels of analyses**: the comparison of the statistical differences between the national average IOp indices to those of the other three analytical levels is important. The computed national level IOp indices are statistically different from those of the eastern, central and western regions, the urban and rural areas, and mostly with all of the different income and consumption groups by quarter.

v. **Between all four levels of analyses**: there are a series of interesting results. a) In general, across all levels of analyses, the computed IOp indices show statistical differences, at least at the 10 percent level. b) The IOp indices from all Development Regions are statistically different to rural areas, and the west and east results are statistically different to the urban areas. c) The IOp indices from all Development Regions are statistically different to the results from the second, third and top quarter earners, but not for the bottom quarter earners. d) The IOp indices from all Development Regions are statistically different to the results of all consumption by population quarters groups at least at the 10 percent level. e) The statistical difference results between the urban rural areas and the consumption/income quarters vary.

3) RF-OLS Regression analyses, correlation matrices and partial IORs

Gender and birth region (Mountain, Hill, Terai) are jointly and separately eliminated from the regressions due to their insignificance and the debate about their inclusion into the analyses at all. Gender does not show much statistical significance across most levels of analyses (RF-OLS), however, its coefficient at the national level is statistically significant and being a woman explains over 15% of opportunity deprivation according to the partial IOR estimates.

R-squared is relatively low for most regression in all analyses at all levels. However, this is consistent with the literature (e.g. exclusion of the effort variable and omitted circumstances). See results interpretation section for more detail.

There is a positive **correlation between mother's and father's education** level and a negative correlation between rural birth region and parental education level, sometimes more for mother's sometimes more for father's education.

Overall, the **partial IOR** estimates (summary of all levels of analyses in Table 3.94) mostly confirm the impacts of the categorical circumstances on outcome and highlight the varying amplitudes across all four levels of analyses and the necessity to disaggregate the relatively homogenous national level results. For instance, while at the national level ethnic group belonging accounts for "only" 14%

of opportunities, for urban areas it is over 34% and in the central region 23%. Furthermore, at the national level, parental education together accounts for 31% of opportunities (similar to the far-west region), while in urban areas it is over 45% and in the eastern region less than 8%.

Concerning the impact of **parental education** on economic outcome, the estimates vary. At the national level and for consumption, father's education level is more important than mother's across all population samples (vice-versa for income, but to a lesser degree). Regional disaggregation, however, shows that across four out of five of the Development Regions (apart from mid-west), father's education level is more important than mother's. In urban compared to rural areas, mother's education levels are more important. These results are partially consistent with the existing literature and partially counterintuitive. The results are reconfirmed by the partial IOR estimates where, for instance, in the central region parental education and especially mother's education are suggested to be important circumstances impacting on opportunity differentials by explaining over 48% of overall outcome inequality. The same trend goes for urban against rural areas, where parental education accounts for over 46% and 12%, respectively, of overall outcome inequality, and mother's education is more important that father's in urban areas while the opposite is the case for rural areas. At the national level, however, there is not much difference between mother's and father's education and combined they make up less than 30% of overall outcome inequality. There are various possible explanations. First, due to male emigration, remaining mothers are over-burdened (generating economic revenue, looking after the family and in-laws, taking care of the children) so mothers take less decision in the HH and the children are freer to make their own choices. Second, however, due to the practice of joint families, other remaining male HH members then become the decision-making or prominently influential members. Third, mothers get to take decisions that involve little resources, but de facto fathers still take all big decisions (regular video calls with the families at home). Fourth, in general urban areas tend to be more liberal due to information availability and population density favouring female and mothers' empowerment. Fifth, however, in Nepal social stratification acts as a within group safety net, meaning that the influence of male within type members still dominates and gives relatively little freedom to women, with mothers education impacting less on children's welfare. Thus the predominance of father's education particularly in rural areas.

So, the parental education impact on education is indecisive in terms of gender variation. Yet, parental education is important across all analyses at all levels.

Caste group belonging takes on the expected signs. Despite that during the author's field research (October 2018) and exchange with individuals from all strata of society the consensus was reached that this social stratification is no longer practiced in Nepal and that much avant-garde legislation is in place to combat caste discrimination, the results of this study portray the prevailing importance of caste group belonging. In brief, there are three main factors that explain the importance of caste group belong in predicting economic welfare for the offspring. 1) The over-burdenment of remaining mothers leads to children being largely brought up by grand-parents and uncles and aunts, i.e. older generations, who tend to be more conservative and traditional. 2) Families that had to remain in Nepal and could not emigrate with the husband tend to be poorer and more conservative. 3) There is a data selection bias for 30 to 49 year olds, who tend to be relatively more conservative than younger generations, and furthermore the NLSS III dataset dates back to 2010/11, meaning that individuals who were born between the 1960s to the 1980s are analysed.

The partial IOR estimates indicate that ethnic minority or dominated caste group belonging explains over 34% of urban area inequality suggesting the circumstance's significance in causing opportunity deprivation. It is the highest partial IOR across all levels of analyses and suggests a continued residential and physical segregation of ethnic minorities and caste groups in urban centres. At the national level, however, ethnic belonging "only" accounts for about 18% of inequality, yet it is the highest share of all categorical circumstances at the national average level.

Buddhists, despite being a minority in the country, afford an elite role in Nepalese society. In Buddhism all are equal and there is no caste stratification. Many untouchables convert to Buddhism and send their children to monasteries so that they can escape poverty and get access to high quality education and discipline. Conversion to Buddhism has slowed down as the Nepalese society evolves from the caste system into the 21st century.

Rural birth region: The national level analyses indicate that there is a negative of impact of being born in a rural area on economic outcome. When disaggregating the IOp analyses, however, the level of IOp in rural areas is lower than in urban areas, and especially the most opportunity deprived groups suffer less from discrimination in rural compared to urban areas.

4) Opportunity deprivation share of circumstances and poverty

At the **national level** and in terms of consumption expenditure, over 4% of the population i.e. 1.14 million people, live below the 1USD/day poverty line. Over 85% of Nepalese live on below 2USD/day. Disaggregating this by Development Region shows that "only" 55% of people in the central region live below the 2USD/day poverty line, while this reaches more than 98% of all people living in the far-western and mid-western regions.

It is evident that **the more circumstances one accumulates** that negatively impact on opportunities, the more an individual is discriminated against and the lower the expected mean per capita income and consumption. In general, for individuals who accumulate 4 or more discriminatory circumstances, there appears to be no difference in gender or having been born in the mountain, hill or Terai ecological belts. There are some variations according to the analysed levels, however. For instance, belonging to the most disadvantaged groups and being born in an urban area in the Terai belt translates into having less than one third consumption expenditure than being born in an urban area in the hill belt. Furthermore, a clear gender gap exists amongst the most opportunity deprived in the mountain belt in the central region. The poorest most deprived women in the centre are born in the mountain belt, while the least poor of the most disadvantaged in the centre, are men in the mountain belt. Their mean level of consumption is more than 250% that of women.

Furthermore, there is a **statistical difference** of mean consumption between groups of people who accumulate different numbers of negatively impacting circumstances **(**no, one, two, three, four or five negatively impacting circumstances). Table 3.92 shows the paired t-test results for the mean differences between groups of people who accumulate different numbers of negatively impacting circumstances. The table illustrates the results using per capita consumption and for the specific circumstantial share breakdowns at the national level analyses. Similar results were obtained for income and for all other three levels of analyses (available upon request). As can be seen from the table, all results are statistically significant at the 1 percent level. In other words, for all pairwise combinations of groups that accumulate different numbers of negative circumstances, the null can be rejected. The means of per capita consumption (and income) from one group compared to another are statistically different.

5) Robustness checks

The **proportion of type distributions** with less than five observations is relatively high, for instance, with 12 to 25% across all levels of analyses. This is due to relatively small number of observations and a relatively large spectrum of circumstantial categorical variables.

Regression tests: At all levels of the analyses, all regressions underwent various tests to check for robustness. These include, for example, Link or Ramsey tests, the residuals were checked for homoscedasticity and normality, extreme values and outliers, etc.

In almost all IOp analyses at all levels, the **correlation matrices** show insignificant levels for gender.

Table 3.91: Statistical Difference Between the Parametric and Non-Parametric IOL and IOR Estimates Across All Levels of Analyses

	IOL (non-para)	**IOR (non-para)**	**IOL (para)**	**IOR (para)**
IOL (non-para)	-	-	-	-
IOR (non-para)	-8.40***	-	-	-
IOL (para)	3.87***	9.42***	-	-
IOR (para)	-3.22***	5.26***	-5.57***	-

Note 1: T-test's (paired) t statistic and statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1. Testing the null that the mean level of the IOL and IOR estimates between samples is equal to 0. *Note 2*: Correlation matrix of the IOp estimates shows that they are all positively correlated and mostly always at the 1 percent significance level. Available upon request

Table 3.92: Statistical Difference Between Means of Types Based on the Number of Accumulated Negative Circumstances, Most Opportunity Deprived (National Level Analyses)

	Number of accumulated negative Circumstances					
Number of accumulated negative Circumstances	0	1	2	3	4	5
0	-	-	-	-	-	-
1	7.92***	-	-	-	-	-
2	13.31***	14.32***	-	-	-	-
3	16.78***	25.26***	18.48***	-	-	-
4	19.04***	32.63***	33.46***	19.38***	-	-
5	19.71***	32.68***	30.27***	17.55***	5.35***	-

Note: T-test's (independent sample, unequal variance) t statistic and statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1. Testing the null that the mean level of *pccons* of 2 of the 6 groups that accumulate negative circumstance variables is equal to 0.

Table 3.93: Results Summary Comparison Table

	National level			Regional Development level					Urban vs. rural		Consumption and income quarters			
	All HH members*	**HHH/ spouse ****	**HHH/ spouse, sons/ daughters, bros/sis*****	**Eastern Region°**	**Central Region°°**	**Western Region°°°**	**Mid-Western Region°°°°**	**Far-Western Region°°°°°**	**Urban Region^**	**Rural Region^^**	**Bottom Quarter Earners & Consu-mers'**	**Second Quarter Earners & Consu-mers"**	**Third Quarter Earners & Consu-mers'"**	**Top Quarter Earners & Consu-mers'""**
Per capita consumption expenditure (variable *pccons*)														
Mean per capita consumption expenditure in NRs. (USD, using *pccons*)	50,466.86 (631.63)	49,712.01 (622.18)	50,640.76 (633.80)	45,027.07 (563.54)	63,777.87 (798.22)	54,893.54 (687.03)	54,893.54 (687.03)	30,092.16 (376.62)	88020.17 (1101.63)	45202.83 (565.74)	16939.96 (212.01)	28302.92 (354.23)	45686.17 (571.79)	110577.30 (1383.95)
Gini coefficient	0.390	0.387	0.391	0.342	0.395	0.416	0.416	0.319	0.414	0.366	0.126	0.074	0.089	0.239
Mean Logarithmic Deviation	0.242	0.244	0.242	0.179	**0.254**	0.267	0.137	**0.184**	0.275	**0.211**	0.028	0.008	0.012	0.096
IOp Index Ratio (non-parametric)	**0.320**	**0.285**	**0.311**	**0.441**	**0.393**	0.285	0.388	**0.441**	**0.442**	0.219	**0.089**	0.124	0.173	**0.180**
IOp Index Ratio (parametric)	**0.263**	0.219	0.250	0.183	**0.365**	0.319	0.171	0.269	**0.343**	0.170	**0.031**	0.032	0.040	**0.072**
Per capita income (variable *pcincome*)														
Mean per capita income in NRs. (USD, using *pcincome*)	110,067.60 (1377.57)	106,818.70 (1336.90)	109,292.50 (1367.87)	102,183.10 (1278.89)	134,464.30 (1682.91)	116,876.50 (1462.78)	68,750.75 (860.46)	74,449.25 (931.78)	175,496.20 (2196.45)	100,932.90 (1263.24)	13353.16 (167.12)	42627.14 (533.51)	87773.59 (1098.54)	280104.30 (3505.69)
Gini coefficient	0.537	0.533	0.535	0.512	0.519	0.545	0.525	0.610	0.483	**0.532**	0.221	0.125	0.123	0.319
Mean Logarithmic Deviation	**0.560**	0.561	0.561	0.484	**0.490**	0.573	0.632	**0.837**	0.464	**0.550**	0.106	0.025	0.023	0.148
IOp Index Ratio (non-parametric)	**0.169**	0.167	0.176	0.213	**0.280**	0.175	**0.318**	0.236	**0.326**	**0.128**	**0.535**	0.077	0.174	**0.226**
IOp Index Ratio (parametric)	**0.109**	0.096	0.109	0.103	**0.188**	0.116	**0.114**	0.091	**0.184**	0.074	**0.017**	0.007	0.037	**0.035**

Note: *For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 5,899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs. ***For all 6,530 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs.

°For all 1,636 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°For all 2,469 obs. for per capita consumption expenditure and 2,464 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°°For all 1,448 obs. for per capita consumption expenditure and 1,423 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.°°°°For all 899 obs. for per capita consumption expenditure and 894 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.°°°°°For all 582 obs. for per capita consumption expenditure and 580 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

^For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.^^For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

'For all 1,859 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. "For all 1,685 obs. for per capita consumption expenditure and 1,1630 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. '"For all 1,723 obs. for per capita consumption expenditure and 1,788 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. '""For all 1,790 obs. for per capita consumption expenditure and 2,033 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.94: Summary of Circumstantial Specific Opportunity Shares (Partial IORs), All Level

	National level			Regional Development level					Urban vs. rural		Consumption and income quarters			
	All HH members *	HHH / spouse **	HHH/ spouse, sons/ daughters, bros/sis***	Eastern Region°	Central Region °°	Western Region°°°	Mid-Western Region °°°°	Far-Western Region °°°°°	Urban Region ^	Rural Region ^^	Bottom Quarter Earners & Consu-mers'	Second Quarter Earners & Consu-mers"	Third Quarter Earners & Consu-mers'"	Top Quarter Earners & Consu-mers""
Per capita consumption expenditure (variable *pccons*)														
Mean Logarithmic Deviation	0.247	0.237	0.242	0.179	0.254	*0.267*	0.137	0.184	0.275	0.211	0.028	0.008	0.012	0.096
IOp Index Ratio (parametric)	0.263	0.219	0.250	0.183	0.365	*0.319*	0.171	0.269	0.343	0.170	0.031	0.032	0.040	0.072
Circumstances														
Ethnic Minority / Dominated Caste	0.140	0.100	0.122	0.040	0.227	*0.701*	0.051	0.121	0.343	0.033	-	0.008	0.017	0.026
Buddhist	0.068	0.024	0.050	0.062	0.139	*0.668*	-	-	0.085	0.034	-	-	-	0.012
Other Religion	0.167	0.137	0.155	-	0.130	*0.666*	-	0.101	0.061	0.078	0.050	-	0.001	-
Father's Education: Primary Complete or Above	0.146	0.106	0.180	0.017	0.219	*0.614*	0.107	0.130	0.206	0.076	0.035	0.015	-	0.037
Mother's Education: Primary Complete or Above	0.155	0.101	0.131	0.060	0.264	*0.661*	0.056	0.128	0.248	0.040	-	0.006	0.020	0.003
Rural Birth Region	0.135	0.090	0.115	0.047	0.225	*0.469*	0.051	0.120	-	-	0.035	0.007	0.014	0.033
Hill Birth Region	0.133	0.092	0.116	-	-	*0.059*	0.016	-	0.105	0.029	0.009	0.011	0.012	0.015
Terai Birth Region	-	-	-	-	0.221	*0.657*	0.051	0.120	-	0.027	0.034	-	-	0.033
Female	0.148	0.112	0.130	-	-	-	0.047	-	-	-	0.034	-	-	-
Per capita income (variable *pcincome*)														
Mean Logarithmic Deviation	*0.560*	0.561	0.561	0.484	0.490	0.573	0.632	0.837	0.464	0.550	0.106	0.025	0.023	0.148
IOp Index Ratio (parametric)	*0.109*	0.096	0.109	0.103	0.188	0.116	0.114	0.091	0.184	0.074	0.017	0.007	0.037	0.035
Circumstances														
Ethnic Minority / Dominated Caste	0.061	0.057	0.063	0.018	0.111	0.272	-	0.026	0.035	0.033	0.032	-	0.012	0.003
Buddhist	0.037	0.033	0.037	0.034	0.062	0.269	-	-	0.025	0.018	0.055	-	0.026	-
Other Religion	0.044	0.042	0.046	0.010	0.071	-	-	0.132	0.001	0.016	-	0.010	-	0.000
Father's Education: Primary Complete or Above	0.070	0.070	0.076	0.048	0.104	0.241	0.096	0.098	0.060	0.060	0.067	-	0.008	0.008
Mother's Education: Primary Complete or Above	0.077	0.062	0.075	0.003	0.135	0.267	0.000	0.098	0.077	0.040	-	-	0.010	0.014
Rural Birth Region	0.065	0.059	0.066	0.002	0.126	0.222	0.008	-	-	-	0.064	-	0.005	-
Hill Birth Region	0.057	0.053	0.058	-	0.111	0.029	-	-	0.003	0.033	-	-	0.019	-
Terai Birth Region	0.063	0.058	0.065	-	-	0.268	-	-	-	-	0.065	-	0.011	-
Female	-	-	-	-	-	-	-	-	-	-	-	-	-	0.010

Note: only circumstances that were statistically significant at least at the 10 percent level were used to compute the partial IORs. *For all 7,044 obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. **For all 5,899 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse and 3) who have information on all circumstance variables out of all survey obs. ***For all 6,530 obs. who 1) are aged between 30 to 49 years 2) are household head or spouse sons or daughters and brothers or sisters and 3) who have information on all circumstance variables out of all survey obs. °For all 1,636 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°For all 2,469 obs. for per capita consumption expenditure and 2,464 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. °°°For all 1,448 obs. for per capita consumption expenditure and 1,423 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.°°°°For all 899 obs. for per capita consumption expenditure and 894 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.°°°°°For all 582 obs. for per capita consumption expenditure and 580 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ^For all 880 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.^^For all 6,174 obs. (both per capita consumption and per capita income) who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. 'For all 1,859 obs. for per capita consumption expenditure and 1,636 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. "For all 1,685 obs. for per capita consumption expenditure and 1,1630 obs. for per capita income obs. who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. '"For all 1,723 obs. for per capita consumption expenditure and 1,788 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs. ""For all 1,790 obs. for per capita consumption expenditure and 2,033 obs. for per capita income who 1) are aged between 30 to 49 years 2) all household members and 3) who have information on all circumstance variables out of all survey obs.

Table 3.95: Statistical Difference Between Means of IOL and IOR Estimates (Parametric and Non-Parametric, for Income and Consumption) Between All Levels of IOp Analyses

	National level*	Eastern Region	Central Region	Western Region	Mid-Western Region	Far-Western Region	Urban	Rural	Bottom Quarter Cons.	Second Quarter Cons.	Third Quarter Cons.	Top Quarter Cons.	Bottom Quarter Earners	Second Quarter Earners	Third Quarter Earners	Top Quarter Earners
National level	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Eastern Region	2.48**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Central Region	-4.64***	-5.08***	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Western Region	-2.67**	-3.55***	4.27***	-	-	-	-	-	-	-	-	-	-	-	-	-
Mid-Western Region	-1.20	-1.89	0.43	-0.40	-	-	-	-	-	-	-	-	-	-	-	-
Far-Western Region	-1.45	-3.20**	1.88*	0.11	0.78	-	-	-	-	-	-	-	-	-	-	-
Urban	-3.97***	-5.00***	1.75	-3.72***	-0.23	-1.52	-	-	-	-	-	-	-	-	-	-
Rural	4.64***	1.85*	4.81***	3.88***	2.44**	3.40**	4.45***	-	-	-	-	-	-	-	-	-
Bottom Quarter Consumers	3.36**	2.61*	3.22**	3.42**	2.44**	2.630*	3.47	3.21**	-	-	-	-	-	-	-	-
Second Quarter Consumers	3.42**	2.81*	3.27**	3.48**	2.57**	2.74*	3.55**	3.03*	-0.25	-	-	-	-	-	-	-
Third Quarter Consumers	3.37**	3.08*	3.29**	3.50**	2.62**	2.85*	3.64**	0.04*	-0.66	-0.89	-	-	-	-	-	-
Top Quarter Consumers	3.44**	2.97*	3.30**	3.56**	2.46*	2.76*	3.70**	2.05	-1.85	-2.41**	-2.98*	-	-	-	-	-
Bottom Quarter Earners	-0.39	-0.55	0.18	-0.31	1.35	0.16	0.04	-0.69	-	-	-	-	-	-	-	-
Second Quarter Earners	6.82***	7.76***	7.03***	7.54***	5.41**	4.61**	5.49**	4.28**	-	-	-	-	1.14	-	-	-
Third Quarter Earners	4.16**	4.60**	10.07***	6.52***	5.41**	4.06**	7.48***	1.44	-	-	-	-	1.04	-1.77	-	-
Top Quarter Earners	2.74*	3.09*	7.70***	6.95***	6.17***	4.79**	7.04***	0.84	-	-	-	-	1.01	-2.14	-1.11	-

Note 1: T-test's (paired) t statistic and statistical significance indications: *** p<0.01, ** p<0.05, * p<0.1. Testing the null that the mean level of the IOL and IOR estimates between samples is equal to 0.
Note 2: Correlation matrix of the IOp estimates between the different levels of analyses is available on request.
*National level IOL and IOR indices include all three national level analysed population samples (three different HH composition sample constraints)

3.9. Is Nepal Trapped? What Can Be Done? Policy Implications and Challenges

This study's results suggest that Nepal may be trapped in terms of socio-economic development and that IOp is a social problem that must be addressed. While the national level results paint a relatively moderate picture of opportunity deprivation compared to the literature, the geographical disaggregation indicates various levels of severance in terms of geographical coverage and the types of people that are the most disadvantaged by their circumstances, that is their environment. This calls for the need to localise and customise opportunity policies that take both a top-down and a bottom-up approach, which should maximise their effectiveness. Further research and impact evaluation analyses should follow beyond this study.

Due to the nature of this study and the dataset utilised, it is important to bear in mind the following. First, the results of this study have a limited country representativeness scope since only a sub-sample of the nationally representative NLSS III dataset was used.

Second, the policy implications are geographically limited due to, for instance, the different cultural, socio-economic or political environment across countries. Also, the analyses here include Nepal-specific elements such as the Development Region disaggregation or the ecological belt and caste circumstance variables.

Third, this study should be complemented with additional research and with a richer dataset (see study limitations and further extensions Section 3.10) in order to draw more precise, pragmatic and effective policy framing to combat IOp.

Fourth, as briefly mentioned before, Section 3.1.6.2 elaborates on omitted variables and the (im)possibility for causal inference. Despite the fact that the lack of being able to make causal inference when calculating IOp, Roemer and Trannoy (2015) emphasise the importance of the correlation and relationship trends that the IOp analysis allows. In other words, being able to say what circumstances correlate with the degree of IOp and compared to the overall level of outcome inequality. Additionally, this study provides a more social stratification specific and geographically fine-tuned analysis allowing for a more thorough understanding of the IOp reality across the country and society.

Overall the scope of this study and its subsequent policy implications are limited in nature. They do, however, provide an important insight into understanding the IOp reality and the socio-economic context of Nepal and can serve as a starting point for a wholistic policy framing package to reform and positively restructure Nepal's long-term socio-economic development.

In order to briefly build on this study's results and the implicit policy recommendations that can be induced, this section takes the following structure:

1) What is the situation? Facts: study results, political institution failure, geographical mobility as opportunity-equalising instrument
2) Prevalence of IOp: international development goals, targets, assistance
3) General policy remarks: short-, mid- and long-term policy planning
4) Four main policy clusters implications: practical approach to combatting IOp

1) What is the situation? Facts: study results, political institution failure, geographical mobility as opportunity-equalising instrument

This study provides information on the most opportunity deprived groups of people in Nepal, the degree of their opportunity deprivation by comparing their mean economic welfare levels to the national means, and the geographical divergence of discrimination across the country. Some of the

results indicate that the share of opportunity deprivation out of overall inequality in economic welfare can be as high as 44% for the most disadvantaged groups (e.g. central region and urban areas). Furthermore, one can conclude that the more negatively impacting circumstances one accumulates, the more and the more disproportionately these will affect the person's economic outcome. For instance, people in urban areas who accumulate four or more of such circumstances are doomed to have a level of consumption expenditure that is only 43% of the urban mean level. This affects over 3% of the urban population. For the central region, over 38% of its population (that is over 345,000 people) accumulate four or more negative circumstances and are guaranteed to live with a mean consumption expenditure level between 39% to 63% of the region's mean.

The results of this study indicate that parental education can explain up to 45% of overall inequality differences and thus also impact on opportunity deprivation according to the partial IOR estimates. Caste group and ethnic belonging are also suggested to explain up to 34% of total inequality particularly in urban areas, for instance. Social and ethnic stigmatisation appear to prevail in Nepal, significantly hindering socio-economic mobility. The results further indicate stark regional differences. The geographical divergence in terms of socio-economic inequality due to caste and ethnic group belonging has also been confirmed by UNDP and the National Planning Commission (2014, Ch. 2 & 3). For instance, they find that when adjusting the Human Development Index (HDI) for caste and ethnic group belonging as well as to gender, that inequality and poverty varies greatly across different caste and ethnic groups and across the country. Furthermore, they find that when the country's different development regions are ranked in terms of access to facilities (such as schools, health care centres), the ranking of these regions has not changed over the past 20 years possibly suggesting minimal socio-economic change and the population's dissatisfaction with the situation.

Nepal has some of the most **progressive legal regulations** to enhance human rights and equality in the region. The Constitution (2015, Preamble) states that

> "embracing multi-caste, multi-lingual, multi-cultural and diverse geographical specificities, by ending discriminations relating to class, caste, region, language, religion and gender discrimination including all forms of racial untouchability, in order to protect and promote unity in diversity, social and cultural solidarity, tolerance and harmonious attitudes, we also express our determination to create an egalitarian society on the basis of the principles of proportional inclusion and participation, to ensure equitable economy, prosperity and social justice".

However, the Nepalese government struggles to successfully execute these. One of the most prominent examples that has gained also international media attention is the treatment of women during their menstruation period (e.g. BBC, 21 January 2019). Nepalese tradition has it that menstruating women are banned outside the house or village to live in separate shelter during this period because they are seen as impure and as omens of bad luck. While this was criminalised in 2017, it still remains a common practice especially in the rural areas and distant areas from the capital like the mid- and far-west regions.

The decade-long failure of the **political institutions** to deliver increased social and economic mobility has forced people to take actions themselves. To complement this study's results, UNDP and the National Planning Commission (2014, p. 65) conclude that "since the 1990s, the question of spatial inequalities and imbalances in development has largely remained unattended". Two additional examples of political failure are highlighted by them. First, the government has been unable to boost the coverage and quality of public social facilities in particular education and health care, leaving out the most disadvantaged and rural areas of equal opportunities. Second, the government has been unable to guarantee a vibrant and expanding economic development in order to provide high quality

employment leading to a mismatch in education, productivity and employment levels. These elements further undermine the legitimacy and credibility of the still fragile and infant democratic Nepalese institutions. This study's results show that therefore, various populations segments of society utilise both domestic and international migration as an **opportunity-equalising instrument.**

First, especially the **most opportunity deprived** are forced to migrate. A) They migrate domestically to urban centres for education purposes. The World Bank (2016, p. 42) highlights that "physical mobility between rural and urban areas is a critical part of intergenerational mobility in education". But as the IOp results at the urban rural area level analyses show, opportunity deprivation areas for the most disadvantaged are amongst the highest in the entire country with over 44% of overall inequality in outcome being due to opportunity deprivation. Since the level of IOp and discrimination is highest in urban centres, once the most disadvantaged receives some basic education, they then emigrate. B) The most disadvantaged migrate internationally as low-skilled labourers (75% of all emigrants, 20-25% of the total population; MLE, 2018) with the hopes to then be able to send remittances to their remaining families. In turn, remittances have become an important contributing factor to Nepal's GDP (ca. 25%; World Bank, 2011), and as a result a substantial pillar of Nepal's economic development strategy. According to Ezemenari and Joshi (2019), real GDP growth is largely fuelled by remittances (through increased domestic consumption) and according to the MLE (2018, Part 3) the Nepalese government has pushed laws, legislation and policies in order to support emigration and guarantee emigrants' fair treatment abroad. In the same report the MLE (2018, p. 53) highlights that "International labour migration is a critical determinant of the employment situation in the country and has been considered a major force of change and development, particularly in poverty reduction and sustainability of economy." Furthermore, these remittances are crucial for the remaining families who have heavily indebted themselves in order to allow for someone in the family to move abroad. It can be argued that a secondary positive effect of these low-skilled labour emigrants is a brain gain upon their return. However, Nepal has entered a vicious downward spiral for economic development with too much emphasis on remittances and international aid, so that when the "newly trained" medium- to high-skill emigrants return, they cannot find employment in Nepal. This is in part because Nepal has not undergone the necessary structural economic reforms and investments in order for the economy and labour market to absorbed more highly skilled returnees. UNDP and the National Planning Commission (2014, Ch. 4 & 5) highlight the mismatch in educational attainments, private sector development and productivity ability in Nepal, where particularly youth employability is extremely low. The World Bank (2016) highlights that Nepal's dependence on remittances has two major adverse effects on its economy. First, the high inflow of remittances has appreciated the real exchange rate, consequently decreasing the competitiveness of the infant manufacturing sectors and further undermining job creation. Second, they argue that the continued outflow of young Nepalese could have disincentivised policymakers to push for and seeing the urgent need for mid- and long-term reforms.

Second, the **more advantaged groups** also migrate. A) They migrate domestically, and especially to the urban centres and the capital to seek higher education. But, the employability of young Nepalese and a serious lack of job opportunities in Nepal make it very difficult for them to make a living UNDP and the National Planning Commission (2014, Ch.4 & 5). This study's results indicate that opportunity deprivation rates are the highest in the urban centres and also amongst the top 25% of consumers and earners. Arguably, Nepal's limited economic development and having a centralised state means that there are more, albeit still few, opportunities in urban areas and especially the capital. However, this also means that it is easier to discriminate based on circumstances, and the greater number of "opportunity privileged" individuals in urban centres are in fierce competition and must leverage all social capital to seize opportunities. In other words, the high population density in urban centres in proportion to the relatively few economic opportunities means that these are reserved for only the most opportunity privileged people. According to Ezemenari and Joshi (2019, p. 47) "Social status determines the types of personal networks individuals can access […] personal networks are an

essential element for youth and adults who seek employment." Individuals who accumulate opportunity depriving circumstances are left with disproportionately fewer opportunities. So, they also have to emigrate eventually (or move back to the countryside). B) The most advantaged groups (about 25% of emigrants; MLE, 2018, Part 4) also emigrate internationally 1) to seek high-quality university education, 2) to seize quality employment, 3) to escape the conservative mind-set and societal pressure at home, such as for out of caste system marriages, and/or 4) for prestige.

To sum up, due to political institutional failure, opportunity deprivation and the lack of economic opportunities in Nepal, both the most disadvantaged and the most advantaged use geographical mobility as an opportunity-equalising instrument. This in turn renders Nepal increasingly dependent on remittances.

2) Prevalence of IOp: **international development goals, targets, assistance**

The topic of inequality and particularly of inequality in outcome has been at the centre of political and public debates for a long time. The coining of the term inequality of opportunity with Roemer's (1998) benchmark book has, however, been a stepping stone towards rethinking the basic concepts of inequality and re-debating the theories of distributive justice (see the literature review sections for inequality in outcome and of IOp). The 2015-launched Sustainable Development Goals (SDGs) by the international community have placed the subject of discrimination and opportunity deprivation at the core for pushing the world's unified development objectives until 2030. Number 10 of the 17 SDG goals focuses on a reduction in inequalities. Sub-target 3 (out of 10) for the reduction of inequalities explicitly highlights the importance of ensuring equal opportunities and end discrimination: "Ensure equal opportunity and reduce inequalities of outcome, including by eliminating discriminatory laws, policies and practices and promoting appropriate legislation, policies and action in this regard" (globalgoals.org[11]). The SDGs do not go far enough for some of the world's leading economists[12], so that they issued what is referred to as the Stockholm Statement. The statement iterates 10 key principles in order to drive long-term economic development globally. Their principle 3 highlights the need for social cohesion and inclusive growth. "Development has to be inclusive. Policy should help ensure that development is socially and economically inclusive and does not leave behind groups of the population – whether identified by gender, ethnicity, or other social indicators." (Stockholm Statement, 2016, principle 3). Stiglitz (2018), who figured on the prominent round of economists for the statement, reiterates that the conservative thought that inequality and economic growth are to some extent mutually exclusive is wrong. In fact, he emphasises that greater inclusivity can lead to more robust growth, and that economic growth and equality act complementarily.

[11] https://www.globalgoals.org/10-reduced-inequalities

[12] Thirteen economists worked on the elaboration of the statement. Professor Sabina Alkire (Oxford), Professor Pranab Bardhan (Berkeley), Professor and former Chief Economist of the World Bank Kaushik Basu (New York), Professor Haroon Bhorat (Cape Town), Professor and former Chief Economist of the World Bank Francois Bourguignon (Paris), Professor Ashwini Deshpande (Delhi), Professor Ravi Kanbur (Ithaca), Professor and former Chief Economist of the World Bank Justin Yifu Lin (Beijing), Professor Kalle Moene (Oslo), Professor Jean-Philippe Platteau (Namur), Professor Jaime Saavedra (Lima), Nobel Laureate Professor and former Chief Economist of the World Bank Joseph Stiglitz (New York), and Professor Finn Tarp (Helsinki and Copenhagen). The meetings were hosted by the Swedish International Development Cooperation Agency (Sida) and the World Bank Stockholm Statement (2016)

Figure 3.21: Stockholm Statement Principle 3: Development Has to Be Inclusive

Source: Stockholm Statement (2016)

As previously stated, Nepal greatly benefits from both financial and technical support from the international community in terms of international aid. The above mentioned international guidelines (amongst many others) serve as an initial point of orientation for policy making and prioritisation. The German Development Cooperation Agency (GIZ), for instance, partners with the Nepalese government on a program called INLCUDE (Inclusive Development of the Economy Programme), which focuses on the socio-economic inclusion of individuals who suffer the most from marginalisation. According to their report and guidelines "development outcomes across all sectors show that gender, caste, ethnicity, geographical location, regional identity and economic status are strong determinants of access to services, resources and political representation" and "exclusion has been a cause and results of unequal development in Nepal" (GIZ, 2013).

As the topic of inclusive growth and equal opportunities gain prevalence also on the international stage, Nepal can hope for increased technical and financial assistance in an effort to effectively combat IOp within its borders.

3) General policy remarks: short-, mid- and long-term policy planning

While the international goals and targets to combat inequality and now also IOp have gained momentum, this study shows that the distribution and severance of IOp varies hugely and has imminent direct effects on the Nepalese population. But as elaborated in the literature review, before coining and prioritising IOp policies a consensus must be reached on the underlying theory of distributive justice between all actors of society, the government, the private sector and the public. Since it is out of the scope of this dissertation to review and analyse Nepal's stance on these, a few general comments and suggestions follow that remain valid for the short-, mid- and long-term effectiveness of IOp policies.

First, Roemer (1998, Ch. 12 & 13) discusses **to what extent opportunities should be equalised**, or in other words, to what extent there is a need to **level playing fields**. He refers to these as Equal Opportunity Policies (EOPs). The EOPs are essentially a set of rules that strive to maximise economic outcome for the most opportunity deprived types. In other words, their goal is to equalise the effect of circumstances on economic outcome, that is levelling playing fields. This would mean

that solely one's efforts, ambitions and decisions are the determining factors for differences in economic welfare amongst individuals; these being in turn arguably more socially and ethically acceptable than opportunity inequalities.

One conventionally agreed on EOP is access to education, where, up to a certain age (or degree level), one can access largely free public education so that everyone accumulates a basic set of tools, which they can use freely to build their futures. This is of course a long-term investment. But the general acceptance of the need for a minimum level of education has given rise to an entire sub-category of IOp literature, which precisely studies the inequality of educational opportunity (e.g. Asadullah and Yalonetzky, 2012; Ferreira and Gignoux, 2014; Schuetz et al., 2005; United Nations ESCAP, 2018). Yet, even the education EOPs are debated and vary hugely. Should education be completely free? Up until what age should education be free? Should this include university level education? Much literature also exists on other domains of EOPs, which are much more debated in terms of the extent and until when to level playing fields. These include, for example, access to health care, the labour market or the credit market.

Furthermore, Roemer acknowledges the difficulty in drawing a clear-cut line between circumstances and efforts, and the extent to which only efforts should predict one's relative success. For instance, should a medical student who tries extremely hard but keeps failing the exams still receive a diploma and be allowed to operate on the sick? Or should a short person who tries hard still be admitted to a national basketball team? So, every society must hold the normative debate and decide to what extent and until when playing fields should be levelled-off.

Second, **personal and public interests** must not be interlinked. Nepal's democratic political institutions are still at an infant and fragile state. The Nepalese public has yet come to understand the meaning of a democratic society, that it is not only privilege and freedom but also duty and compromise. In order to debate the previously discussed levelling playing field policies, it is crucial to overcome the link between personal and political interest. This goes along the line of "it is easier to give a man a fish, than it is to train him how to fish". For political interests, it may be easier to use short-term redistributive measures rather than long-term educational investments in order to level playing fields.

Third, there is a need to figure out how much of the exogenous **circumstances can actually be altered and how "efforts" can be incentivised**. For instance, what factors drive people to enhance their efforts? Does the meaning (i.e. fulfilment), the economic remuneration (i.e. well paid) or the security (i.e. permanent contract) encourage people to enhance their efforts? For sure, the underlying basis is the development of a diverse and dynamic economic system. Furthermore, as birth region appears to affect economic welfare, to what extent can this be altered, or should other, for example rural area development programs take priority to help alleviate the significance of being born in a rural area. Would it make sense to scrap caste group belonging recording in administrative documentation, or should one just ban ethnic and caste group discrimination?

Fourth, no matter how progressive the laws and how successful their execution, **behaviours and customs must also change** (interview with Ms Silwal, KOPILA Nepal, 14 October 2018). This is the most difficult element to change, and it takes time. Arguably, long-term sustainable development, education and an increasing level of economic welfare should bring this about in the (very) long term.

Fifth, the results indicate that Nepal and especially specific population sub-groups are trapped and suffer significantly from socio-economic stigmatisation. It may be suggested that in order to graduate from this status quo or "lock-in" stage, there is the need for a **shock or a major "disruptive" event** (David, 2007). This argument follows the path dependence logic in the literature of the historical

institutionalism approach, which has started to gain momentum in the economics literature on the middle income trap.

Last, unless the government manages to **guarantee a minimum level of socio-economic welfare** for the entire population and can successfully afford to serve as such a safety net, the existing social groupings (i.e. caste, ethnic) will continue to act as socio-economic safety nets and discrimination will prevail. According to Ezemenari and Joshi (2019, p. 47) "Social status determines the types of personal networks individuals can access." Long term inclusive and sustainable economic development should enhance the ability of the government to provide social security so that individuals and families need less rely on social and income group networks as insurance mechanism.

4) Four main policy cluster implications: practical approach to combatting IOp

Only four main clusters of brief policy implications are drawn on here. The first focuses on the overall opportunities Nepal's new federalisation process bears. The following three focus on the most opportunity depriving circumstances, which this study's results revealed.

First, this study's results clearly show that a **"one size fits all" policy** approach would not be suitable. The disaggregation of the IOp analyses allowed for pinpointing variations in opportunity deprivation according to geographical and socio-economic factors. Profiles for the most opportunity deprived types of people and individuals were isolated. Sometimes a gender gap or even a difference in ecological belt origin are observed for the most disadvantaged. This study's results in important regional and differences in persisting inequality and IOp across the country have been confirmed by other literature (e.g. Ezemenari and Joshi, 2019; Kruse et al., 2013; UNDP and the National Planning Commission, 2014). EOPs, given the Nepalese context, must thus be customised and localised. The currently underway political restructuralisation process towards a federal state bears great opportunities and challenges to enhance EOP effectiveness. Some scarce and inconclusive literature exists on the effectiveness of political decentralisation on, for instance, local health, welfare and education improvements or poverty reduction (e.g. Abrigo et al., 2017; Abrigo et al., 2018; Cavalieri and Ferrante, 2016; De Novi et al., 2019; Kelekar and Llanto, 2013; Kis-Katos and Sjahrir, 2017; Jimenez-Rubio and Garcia-Gomez, 2017). Cavalieri and Ferrante (2016), for instance, found that when the mutual responsibility between the local authorities and its people are enhances (e.g. through fiscal payments going to local rather than central authorities), then infant mortality rates drop. They also find that benefits from particularly fiscal decentralisation depends largely on the wealth of the regions, and most so for the most impoverished regions. Jimenez-Rubio and Garcia-Gomez (2017), however, find exact contradicting results. According to them, long-term positive effects of "full decentralisation" (both political and fiscal powers) only prove effective in the wealthiest regions.

Overall, one can conclude from the decentralisation literature that in order for decentralisation to be effective for and benefitting the local people, the process needs to be smooth, with complete integration and training of the local populations and authorities. Furthermore, a maximum effort must be made to achieve complete transparency for public spending and political nominations and elections. A major challenge, however, is that the Nepalese government is still undecided what parts of the central government (e.g. fiscal, administrative, political, legislative) and how much of its sovereignty will be transferred to the local authorities. Many more details and decisions have yet to be taken despite years of trying to push forward with the federalisation process. Without a clear roadmap of how to execute the decentralization, and the necessary trust and political stability to implement this, a major threat is that Nepal's government remains in a status quo situation with bound hands and inability to act. In fact, the reallocation of central budget to local authorities has seen many delays and oppositions so that local authorities have increasingly faced budgetary shortcomings, not being able to provide public services This has disproportionately affected the most remote rural and impoverished areas. Ezemenari and Joshi (2019) highlight some of these challenges and display some of the policies that were passed in a desperate effort to try to unblock and reallocate public budget. For now, it is difficult to draw conclusions about the definitive decision taken for the federalisation process and the effects these will have on the livelihoods of the local people. Ezemenari and Joshi (2019) suggest for a smooth, swift, consistent and coordinated implementation procedure for the federalization process, with particular focus on those that are ought to boost investment in human and physical capital in order to uplift inclusive and sustainable growth.

A particular strength should be the increased political leverage and voice of administrative sub-regions being able to reflect the realities and interests of the locals (Ezemenari and Joshi, 2019, p. 7). Also, giving them more power, autonomy and budgetary means should boost their effectiveness to employ their local networks and know-how. The argument here is that the spending locally raised taxes should harness great accountability of the institutions towards their local citizens. Furthermore, training for increased awareness of IOp at the location administrative level could prove crucial and a clear roadmap plan would need to be developed.

Second, the results highlight the importance of enhancing basic **education levels** and those for all, no matter the gender, ethnic or caste belonging, religious belief or birth region of the

individual. Complementarily to this study's results and the importance that parental education has on children's economic welfare, the CIA World Factbook (2010) has ranked Nepal 190th out of 206 countries in terms of literacy. Despite the great efforts of the Nepalese government and partnering international aid institutions to improve access to public education, much remains to be done. The government must invest into a more inclusive education system with higher quality and greater coverage, particularly in the rural areas. The World Bank, UNDP, NGOs and academics have published extensively stressing the importance for educational investment (e.g. from primary and basic education to high level education that matches the needs of the economy). Ezemenari and Joshi (2019), for instance, highlight in their Human Capital Index (HCI) analysis of Nepal that the access (i.e. coverage) and quality of all levels of education needs to be drastically improved. This is particularly crucial for addressing geographical socio-economic disparities across the country. Furthermore, as the level of primary education increases, the secondary schooling sector needs substantial investment, especially in the rural areas. This is to a) increase the education level of the most disadvantaged, and b) disincentivise them to migrate to urban areas (hubs for better education and work opportunities), where actually they are disproportionally more discriminated against based on their circumstances. Ezemenari and Joshi (2019, p. 14) point out key education investment areas. First, the need to invest into early childhood education programs (e.g. increase in quality and coverage of early childhood education, local authority monitoring capacities, assessment to evaluate child grants to strengthen household incentives for investing into human capital). Second, they stress the importance for specific labour market programs in an attempt to match the needs of local economies (e.g. skill training, self-employed coaching, training for women and youths). Third, invest into high level education (e.g. through performance based grants, standardised exam system to include cognitive and critical thinking). Furthermore, teacher training, quality and certification needs to be standardized at the national level needs to be implemented as well as developing a scheme for emigrant returnees to teach about tangible experiences and skills learnt while abroad (personal interviews, Ms Yamamoto, Cheer Japan Overseas Cooperation Volunteers, October 2018). Also, school graduates need training to prepare them for the labour market, because they do not know what awaits them and what skills are needed. This is especially the case in remote rural areas.

There are a number of challenges in trying to operationalise the education policy recommendation above. First, the lack of public budget and mutual responsibility between the citizens and the state renders educational investments difficult. This is aggravated by Nepal's dependence on remittances, which also leads to brain drain, further disincentivising the government to invest into education, particularly into secondary and higher education. Second, Nepal's appalling development status has increasingly made it dependent on international aid. Given that Nepal has limited resources, it is lowering its priorities for education investments, which is a particular focus to the international community. Third, this is aggravated by the fact that the increased inflow of remittances has led to the development of a dual education system (public and private), further disincentivising the government to spend money on free public education. Suvedi et al. (2014) elaborate in length on the changes in Nepal's education system and the impacts of its privatisation. According to them, over 15% of primary and 32% of secondary schools are private, and the share is growing rapidly. Over the past 6 years the share of private schools grew by over 5% every year with governmental spending for education stagnating (Ghimire, 2018). Currently less than 10% of the national budget (globally, average commitments are around 20%; Ghimire, 2018). So, the reality paints a daunting picture. Maybe there is need for a disruptive event to reverse these happenings, as David (2007) might emphasise.

Third, this study's results show that a) the most opportunity deprived actually suffer less in the rural than in urban areas. However, b) being born in a rural areas greatly aggravates individuals' likelihood to increase their economic welfare. This suggests the need to improve the economic **development of the rural areas**. The Nepalese government began to study the benefits of Special Economic Zones (SEZ) across the country from 2000. By the end of 2000, the first zone was

established and until as of 2017, there are 14 SEZ across Nepal (Special Economic Zone Development Committee, 2017). However, their relative effectiveness in boosting local economic spill-overs and increasing welfare standards for the communities remains to be studied. A key challenge in their operationalisation is the governments strategic pillar and dependence on remittances, at least in the short- and mid-term. This has and continues to undermine industrial and service sector development, which is a prerequisite for creating more employment opportunities for people who would otherwise emigrate or who return to the country. Yet, Nepal has passed a series of policies to attract Foreign Direct Investment (FDI) and help boost the potential success of the SEZs. These policies include the Public-Private Partnership and Investment Act, the Foreign Investment and Technology Transfer Act and the Transit and Transport Agreement with China. Ezemenari and Joshi (2019, Chapter B) estimate that these should substantially boost FDI into Nepal, stimulate GDP growth and drive private investment and consumption.

Fourth, promoting and **executing minority and human rights** across Nepal is a crucial element in alleviating opportunity deprivation and discrimination. Nepal struggles to successfully execute its relatively progressive human rights laws. Women, for instance, are still regarded as impure during menstruation, particularly in the countryside, and people from the untouchable caste communities are refused access to the credit market or the ability to defend their rights in the existing justice system. The government has, however, implemented new representative quotas for women and minority groups in its public institutions, which lie under the functions and duties of a newly structured institution called the National Inclusion Commission (Constitution of Nepal, 2015, Article 259). The current federalisation process can be seen as an opportunity to give the marginalised a voice, since the real discrimination and social exclusion challenges lie within the communities (Interview, Ms Silwal, KOPILA Nepal, 14 October 2019). The challenge, however, is the meaningful and not purely physical representation of women and minority groups. Often they cannot speak honestly of their socio-economic backgrounds because they risk losing their jobs. Furthermore, there is a conflict of interest within these groups, where it is in their interest to guard the status of a discriminated minority group.

Bhattachan, who is a Nepalese sociologist and minority activist, extensively describes minority issues in Nepal in his 2003 "Indigenous Nationalities & Minorities in Nepal" report and gives concrete policy recommendations. According to Bhattachan (2003, p.44-48), Nepal must find a way to political stability if it wants to prevent further violence and hatred amongst and towards minority groups. "If democracy should prevail, pluralism or diversity should be respected, minorities would have an opportunity to have a meaningful participation at national to local levels. (…) If all concerned authorities and peoples take socio-cultural diversity as strength Nepal has a bright future." Some of his recommendations include the elimination of all forms of discrimination against all minority groups, the creation of a secular state and the implementation of a minority-led development strategy.

This section demonstrated that Nepal must take EOPs seriously if it wants to achieve socially just and prosperous economic development. Nepal has limited public resources to do so, but a pragmatic bottom-up and top-down approach with localised policies is needed for effective policy implementation. Despite some of the steps that Nepal has already taken to combat rural degradation, minority discrimination and illiteracy, this study's results demonstrate that significantly more dedication, social change and political will must be given in order to effectively level playing fields and incentivise efforts.

The policy implications indicate that rural development schemes, the development of a solid federal state with political and budgetary transparency, a strong and executable jurisdictional framework for social tolerance and inclusion should be given priority. Also, enhancing the quality of and access to free public education is not only key to boosting literacy and employability rates but also

to facilitating change in discriminatory behaviours and strengthening the societal roots of Nepal's infant democracy.

3.10. Limitations and Further Extensions

This section first briefly elaborates on the various limitations of this study, including data constraints, methodological shortcomings and non-generalisability of the policy implications. Second, some further research extension ideas for the IOp analyses of this study from a theoretical and an empirical standpoint are given.

3.10.1. Limitations to the Study of Inequality of Opportunity

The **methodological limitations** of this study's IOp analyses are threefold. First, the exclusion of the effort variable can be seen as a shortcoming to the completeness of this IOp study. The methodology used in this study does not include the effort variable despite it being considered an important element for explaining economic outcome and being a crucial element of the IOp analyses. This was mainly for two reasons. A) The need to accept the independence assumption between the effort and circumstance variables, and B) the difficulty to accurately measure effort. Section 3.1.4 elaborated on the challenges and trade-off of the IOp analysis from a more ideological perspective, which was complemented by Section 3.1.6 further extended the debate on the operationalisation possibility of the effort variable.

Second, there are a number of unobserved variables that were not included in this study. Some of this was due to the nature of the dataset (e.g. the nonexistence of information on parental occupation), some of this was due to methodological trade-offs (e.g. trade-off between the maximum inclusion of circumstances and low type distribution; e.g. more categorical variable or additional circumstances such as HH composition or presences of parents) and some of it is due to intrinsic nature of certain variables (e.g. luck or genetics). Yet, this study included 7 circumstance variables, more than, for instance, FG (2011) which served as the key reference literature for this study's analyses.

Third, this study is a pure cross-sectional analysis of the IOp reality of Nepal in 2010/11. The following section proposes a possible panel analyses extension, but mainly from a theoretical standpoint. This is because the operationalisation of such a model is a) very complex, and b) requires extremely rich panel data, which is not available for Nepal. To the knowledge of the author, empirical panel IOp analyses have yet to be done. Most likely, the first studies will be done for developed countries that have richer datasets over a larger time span.

Two main dataset limitations are highlighted here. First, a richer and more exhaustive dataset would allow for a more precise IOp study. Despite the NLSS III dataset being relatively rich for a developing country Living Standard and Measurement Survey, more information on socio-economic indicators would have been desirable for a more complete IOp analyses. For example, parental occupation, which is one key variable in many empirical IOp studies, does not figure in the NLSS III dataset. Despite various efforts to find a suitable proxy the variable (or a proxy for it) could not be included decreasing the IOp estimators' accuracy. Furthermore, the lack in detail of the parental education variable obliged the author to combine both literacy and education level elements for the parental education circumstantial variables. Despite literacy and education attainment being two separate concepts of measuring education levels, the NLSS III questionnaire combined these two, where respondents could reply with either or. No explanation was given in the NLSS III reports (CBS, 2011) as to why respondents were given this choice. A more detailed fragmentation of the variable would allow to distinguish more precisely between the impacts of the illiteracy-literacy and the

educational attainment benchmarks and their respective impacts on opportunity deprivation. This would be of particular interest given the Nepalese context where education and literacy rates are still extremely low. The author was told by officials from Nepal's Central Bureau of Statistics that the next round of the NLSS dataset (NLSS IV) will have more exhaustive information on, for instance, HH roaster and HH member cultural and social elements. The NLSS IV dataset was collected in 2017/18 and is supposed to be made available for purchase by the end of 2019.

Second, a larger and more representative dataset would render the econometric IOp tests and methodology more reliable and robust. This is especially important to the geographically disaggregated IOp analyses into the Development Regions and urban rural areas, which try to increase the picture clarity of the opportunity deprivation profiles across the country. This, in turn, should allow for a more effective and customised policies.

Last, due to the data and methodological limitations of this study, also its **policy implications** are limited in nature and scope. In order words, a more complete and further disaggregated IOp analyses (e.g. of within caste group and between age cohorts) would allow for more precise policy recommendations. Also, the highlighted recommendations are not generalisable to other countries or regions since the analyses here include country specific variables (such as country specific caste grouping) and the cultural, socio-economic and political realities are distinct for Nepal.

3.10.2. Further Research Possibilities

There are countless ways in which this study could be extended theoretically and empirically, horizontally and vertically. This sub-section will first elaborate on a number of more general possible research extensions and then go into more details of five precise examples.

Limited literature exists that develops a theoretical **panel analysis** IOp model (e.g. Salvi, 2007; Roemer and Ünveren, 2012; Roemer and Trannoy, 2015). Moreover, the empirical application of this is very difficult and has thus not been done much. One reason is that it is difficult to distinguish the value-added of a dynamic IOp analysis approach to an intergenerational mobility approach. The NLSS III dataset used in this study includes a panel sample of 1,032 HHs, of which 513 HHs were also tracked and enumerated in NLSS II and 519 HHs were visited during both the NLSS I and II. Unfortunately, these panel observations are not nationally representative. This implies that, for a thorough, reliable and robust panel IOp analyses, this dataset is not suitable. Yet, all seven circumstance variables that were used in this study are included in all three NLSS rounds. Nevertheless, a cross-sectional comparison between the three rounds could be done.

From a vertical perspective, it could be of value added to **further disaggregate** the analysis. The first example would be an age-cohort breakdown in order to check for an age effect (Hassine, 2011). One would expect that the younger an individual, the higher the impact of his exogenous environment. This is because one might expect circumstances to dominate in earlier life stages, while efforts gain in importance towards the latter life stages, equalising circumstances. A second example would be a within caste group disaggregation. It is commonly known that, for example, Brahman, Chhettri or even the Dalit communities follow a strict within group hierarchy (e.g. Bennett et al., 2008; Levine, 1987; Höfer, 1979).

Nepal is characterised by heavy emigration (over 20% to 25% of its population currently living abroad; MLE, 2018) and substantial economic dependence on remittances (about 25% of GDP; ADB, 2018). The NLSS III survey contains some information on absentees (i.e. labour emigrants that used

to belong to the HHs and have been absent for more than 6 months). An interesting further analysis to this study would be an investigation to what extent a) the **absence of a father (or parent)** and b) **remittances impact** on the level of IOp. Vogel and Korinek (2012), for instance, investigate the impact of remittances on the allocation of children's education expenditure. A further disaggregation to see whether there are differences in allocations across siblings could also contribute to the in-depth understanding of intra-household resource allocation and opportunity deprivation.

Furthermore, an **event impact analysis** could contribute to the literature of IOp by possibly combining it with the path dependence approach discussed by David (2007). Such possible evens could include the industrial structural changes in the US in the 1990s (IT growth and decline in manufacturing) and 2000s (restructuring of the investment banking sector), the time before and after Apartheid in South Africa, or the fall of the Berlin Wall in East Germany. However, much of this subject to data availability.

Further value-added to this study could be the complementary utilisation and comparison to **other IOp measures.** This study's methodology is largely based on the direct ex-ante approach by FG (2011). A comparison of these results to those obtained by the linear regression model approach Bourguignon et al. (2007, see literature review section) or the IOp indices approach utilised by Asadullah and Yalonetzky (2012) would allow to check for further result robustness and to draw better policy implications. The author has not yet come across a paper that empirically compares various IOp methodologies.

A circumstantial variable that the author has not yet come across in any of the IOp literature is **security**. This element could have two circumstances. *Human security* would be a measure at the personal level (i.e. including crime rate elements) and *national security* (i.e. threat of territorial, digital or economic insurgence) would be at the national level. First, a theoretical framework of quantifying these two security measures would need to be developed. Then, adequate data with a divergence across the population would need to be collected and analysed. The national security element could serve in either a panel or a cross-country comparison.

From a horizontal expansion viewpoint, it would be interesting to cross-compare this study's results to **other caste societies** like India, Pakistan or Sri Lanka, and **"caste-like" societies** such as Japan or South Korea (De Vos, G. and H. Wagamatsu, 1966).

3.10.2.1. The Case of Market Liberalisation and Inequality of Opportunity

From an aggregate economics standpoint, it would be interesting to extend the IOp analysis to other micro- and macroeconomic elements such as business innovation, entrepreneurship, or GDP growth. Checci and Peragine (2010), for example, attempt to find out how a change in GDP growth impacts on IOp and vice versa. Furthermore, the study of the impact of market liberalisation on IOp (e.g. utilising the East and West Germany case) could be of relevance to a Nepal that is trying to open up its economy (e.g. through SEZs) and diversify its economic system. But this could also be relevant to understanding the challenges the world faces today.

The underlying methodological approach would be a combination of the IOp methodology and the classical Heckscher-Ohlin model (Bourguignon et al., 2007; Bourguignon and Morrisson, 2002; Fields, 2004). This model looks into different factors of production endowment and assumes the free movement of capital and labour. In theory, trade liberalisation should decrease prices of goods and services, enhance competitivity, productivity and innovations and increase complementarity, wages, production functions and so on. So, theoretically wage inequality should decrease with trade

liberalisation (in both developed and developing countries). The empirical literature, however, paints a slightly less positive picture (e.g. Bourguignon et al., 2007). Instead of doing the habitual liberalisation-inequality analysis one could combine this with the IOp methodology. A disaggregation of, for example, highly and low skilled labourers could be done. However, both a theoretical model and its operationalisation have yet to be developed. Further one could attempt to sub-analyse the impact on e.g. high-skilled and low-skilled labour (Williamson, 1998). Also, a comparison could be done between individuals who are at the peak of their career, before and/or after. The timing of liberalisation can be considered as exogenous.

Nepal ascended as a member state to the World Trade Organisation (WTO) in 2004, i.e. the timing of the second NLSS data round. So, possibly the NLSS datasets could be combined with available trade statistics. It would more likely be a cross-section, than a panel analysis, however, due to the lack of observed years where usually more than 20 to 30 years are needed to measure the real impact of trade liberalisation (Williamson, 1998).

3.10.2.2. The Case of Female Household Headship

The results of this study revealed that there is no consensus on the importance of mother's education. Sometimes mother's education levels appear as more important determinants of children's economic welfare and opportunities and sometimes father's education levels appear to have more influence. It would therefore be interesting to dig further into this topic using the IOp concept and including a more precise gender disparity, motherhood and female HH headship breakdown. Much literature suggest the importance of mother's education and headship for the welfare of children (e.g. Fuwa, 2000; Handa, 1994; Chant, 1997; Buvinic and Gupta, 1997). However, the results of these studies are nonconclusive, and somewhat counterintuitive. Fuwa (2000) highlights the importance to clearly define HH headship, which in reality is difficult to do. Furthermore, much of the available data is subject to the subjectivity of the respondents, who self-report their own HH status. This is all the more troublesome in the Nepalese context where a significant number of husbands are absent (migration). Fuwa (2000) can serve as a reference on the definition of female HH headship. However, the extent to which this can be incorporated into the HH level IOp analysis is subject to debate.

3.10.2.3. Surname Analysis for Social Mobility: The Scholl Family

This extension idea is out of the scope of a thorough largescale microdata analysis, but could serve as a complementary qualitative case study approach. It provides a) an alternative for gathering (qualitative) data to analyse IOp and add depth to the here presented quantitative IOp analysis with a case study, and b) an illustration of the influence of circumstances on efforts.

a) Clark (2014) in his book "The Son Also Rises: Surnames and The History of Social Mobility" elaborates on an alternative methodological approach to studying intergenerational mobility. Instead of using microlevel data, he expands on family surname analogy and professional background as circumstantial condition. Höfer (1979) thoroughly investigates into the origins of caste grouping, belongings, origins, hierarchical structures and so forth. Many caste groups, but not all, having direct associations to certain professions. A theoretical and empirical development of Clark's approach combined with this study's IOp methodology could prove interesting. But first, being based on Labour Force Surveys, the NLSS surveys might not be exhaustive enough on socio-economic and anthropological elements to undertake a social mobility study in terms of Clark combined with the concept of IOp. Second, the NLSS surveys do not span over a large enough time frame i.e. across generations to thoroughly study shift in intergenerational social and professional mobility. If it were possible to obtain a full family tree line of one or several Nepalese families, then Clark's qualitative approach of studying intergenerational social mobility could be used to complement the quantitative

IOp analysis here. The approach could serve as an alternative source of (qualitative) data and complement the IOp analysis as a case study, which in turn could bring value-added to the IOp analysis here.

b) The suggestion made here is an expansion of Clark's family surname analogy methodology to the field of IOp using the case study of the Scholl family, a German family. While the theoretical model has yet to be developed, some purely descriptive investigation was done to the Scholl family tree. In sum, while marriage was geographically mobile, usually marriages were within the same socio-economic strata. The sole up-marriage examples are when wealthier and more influential women were married. The Scholl family illustrates the linkage between efforts and circumstances, and that immobility does not necessarily have to be bad when the immobility is seen as guaranteeing an economically acceptable situation.

Figure 3.22 shows a part of the family tree[13] that remounts back to 1484 and descends all the way down to the author, Sebastian Scholl. When looking at the entire family tree it can be seen that there were four key families. Each family holds and conserves a specific category of profession that each require a set of distinct skills. Figure 3.23 summarises those four main families and their four strands of professions that they guarded for over 17 generations, that is over six centuries. There was almost complete intergenerational (occupational) immobility for over 17 generations, until the 18th generation, the generation of the author. The 18th generation who does not follow into any of the four families' footsteps from a professional perspective.

The question lies in, what did it take for each generation to just accept their circumstances and professional heritage? What factors affect our perception of our self-potential? In other words, to what extent do our circumstances make us believe in our reality, or again differently put, to what extent are our efforts really independent of our circumstances? Looking at the Scholl family tree, one could argue that each of the four families experienced complete intergenerational occupational immobility. Yet each of the families was extremely successful with what they were doing and constantly being able to re-innovate their professions, adjusting them to become, respectively, pioneers of their time in Germany (and Europe) at the time. The Haisch craftsmen family, for instance, having been one of Germany's wealthiest families at the time, made their success in furnishing premanufactured wood from the Black Forest (southern Germany) to the Dutch and Brits up the river Rhine, whose thirst to colonise the world by sea was untamed. Or the Luz Scholl family who ran one of Europe's most successful luxury hotel companies that served as wellness refuge and meeting point for the European and world elite at the time. It is interesting to note, that the only ones who seized their change, and tried to live the "American Dream" by emigrating to the US in the 1920's were from the farming families. It is also interesting to note that many of the American Fortune 500 families that rose during the American Dream area, no longer exists (e.g. Rockefeller), while many of the big European and German families still exist and pull influential strings (e.g. Welfen family). Why is that?

So, what is it that made the pre-inscribed trajectories last for over 17 generations? And that the 18th generation falls out of the line? In other words, despite the children and children's children having had the economic chance to peruse different professions to those that traditionally ran in the family, they chose not to do so (very few exceptions)? Furthermore, when we debate IOp and intergenerational mobility, we usually centre arguments about the discriminated and how to ameliorate their upward mobility and chances to succeed. However, the same debate disappears when already affluent individuals or families conserve their status. As long as their chances do not discriminate or hinder that of the less well off, then their influential positions should not be subjected to inequality discrimination.

[13] The family tree illustration spans over several pages, only the most relevant section is shown here.

Overall, Clark's family name analogy methodology could a) provide additional qualitative data in understanding IOp and thus add additional robustness and understanding to a localised IOp study. b) The brief case study provided here of the Scholl family tree clearly illustrates the link between circumstances and efforts, and raises a number of (ideological) questions on when to begin to equalize opportunities or combat social immobility, and if this should only be an issue of the most discriminated.

Figure 3.22: The Scholl Family Tree (1484 - 2019)

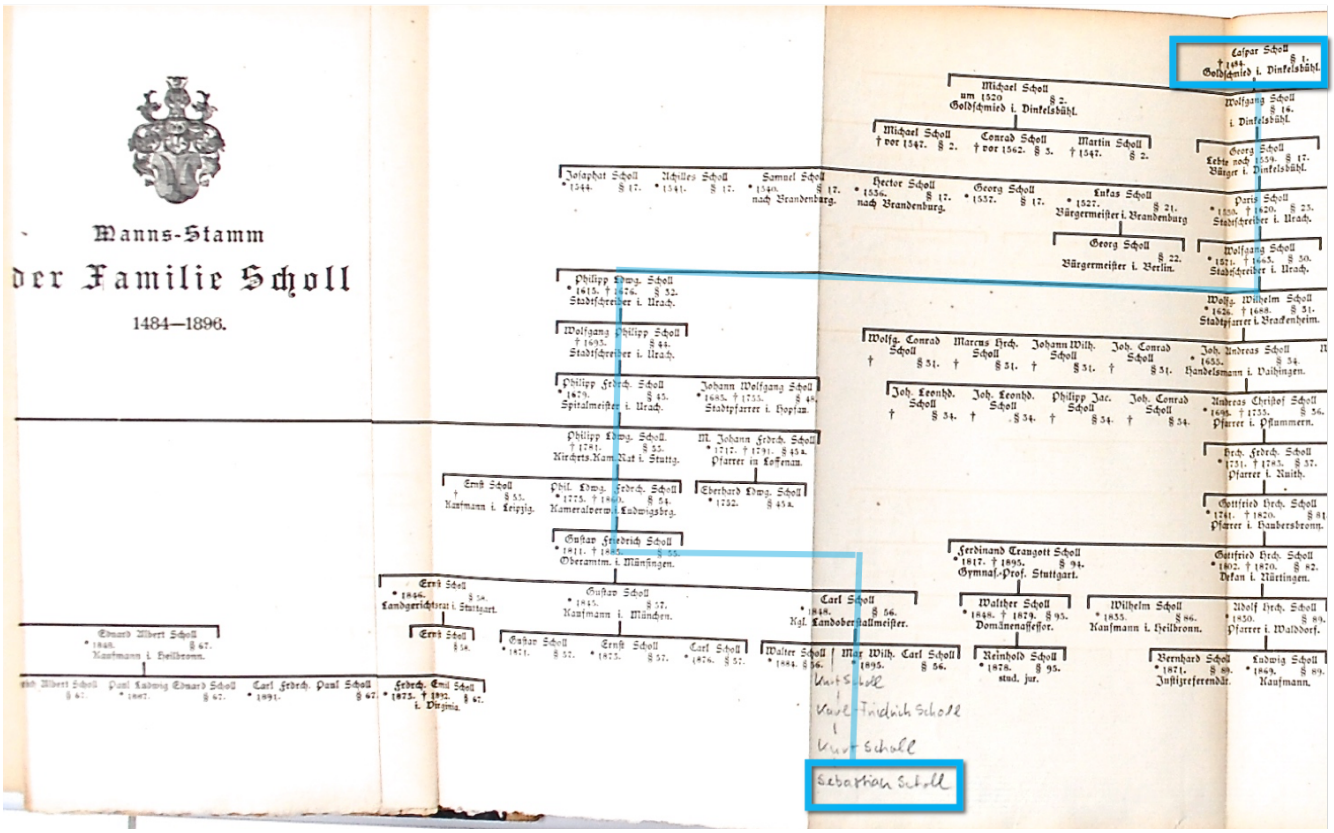


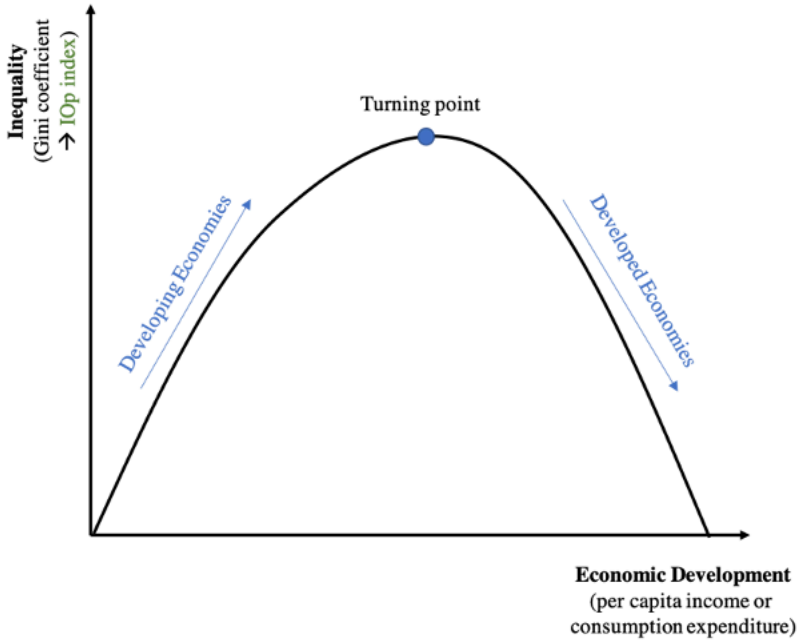
Figure 3.23: Scholl Family: 4 Strands of Professions for Over 17 Generations

3.10.2.4. The Kuznets Curve Revisited.

This sections aims to extend the scope of this study to an empirical country comparison by revisiting the Kuznets Curve (Kuznets, 1955; Fields, 1980 and 2001). The theory and underlying logic is that, as countries develop, inequality in outcome increases up until a certain turning point where it decreases again. This constructs the inverted U-shaped curve illustrated in Figure 3.24 where the Gini coefficient is the chosen inequality in outcome index on the y-axis and the countries' economic development level on the X-axis. The basis of this being the trickle-down effect of development, where structural change comes with development, which over time, should help to alleviate inequality. The empirics of the Kuznets Curve, however, is debated (Fields, 1980; Deininger and Squire, 1988). The proposition here is to use the computed opportunity index here instead of the Gini coefficient on the y-axis.

The IOp analyses of Nepal and their comparison to the empirical results in the literature show some first indications for this. Nepal ranks as one of the 30 poorest and underdeveloped countries in the world (WDI, 2017; The World Bank 2016) and its IOp results are relatively low compared to other slightly more developed, but yet emerging countries such as Guatemala, Panama, Brazil, Egypt or Tanzania (GF, 2011; Hassine, 2011; Hassine and Zeufack, 2015). Developed countries such as Italy, Germany or Norway (Checci and Peragine, 2010; Molinas, et al., 2010), however, face much lower IOp rates than those of the emerging markets. So, the yet to be econometrically tested hypothesis is that, initially the level of IOp rise with development until it reaches a peak point from which on the country continues to develop but sees a decrease in IOp levels. If this hypothesis holds, the next imminent questions are: what does the future, in terms of IOp development, bear for Nepal and to what extent should EOPs be given priority?

Figure 3.24: The Kuznets Curve Revisited



Source: Author's illustration

3.11. Conclusion of the Study: Inequality of Opportunity in Nepal

This section concludes the chapter on IOp, a concept that was only conceived as such relatively recently with the benchmark book by Roemer (1998) "Equality of Opportunity". This chapter highlighted and compared key reference theoretical and empirical literature. The importance of the study of the study of IOp for the public, policy makers and economists was stressed by the shift in public opinions about inequalities in outcome towards IOp. The concept of IOp finds its basis in the argument that inequalities in outcome that arise due to difference in efforts and taken decisions for which one can be held accountable and responsible are more just and socially acceptable than inequalities that arise due to differences in opportunities. Yet, there are numerous challenges and trade-offs with the IOp concept initially developed by Roemer and then further extended by others. One main difficulty being the challenge to operationalise the IOp theory, particularly through the difficulty in measuring efforts.

This study is mainly an empirical analysis trying to understand the current IOp situation in Nepal, rather than being a theoretical investigation. To do so, this study replicates the model developed by FG (2011) with some methodological and variable adaptations for it to suit the NLSS III dataset and fit into the Nepalese socio-economic, political and cultural context. This study also excludes the heavily debated effort variable from the IOp analyses. In fact, much of the empirical literature excludes efforts, because the inclusion of it requires the necessary independence assumption between circumstances and efforts, which is arguably seen as too strong.

The Nepalese NLSS III dataset was tested for IOp and opportunity deprivation profiles of individuals were singled out and cross-compared for four different analytical levels. The first is the national level analysis, which was then disaggregated into Nepal's Development Regions. A further

analysis disaggregated the national level into urban and rural areas. At last, the level of IOp was analysed where the population was divided into quarters by their economic welfare levels.

Despite the national level results being somewhat low compared to the ones of other developing countries in the literature, the results were still signalling. However, particularly the geographical disaggregation indicates that there are striking divergences across the country in terms of opportunity deprivation, and in terms of types of people, i.e. the most opportunity deprived and the most advantages. It is this disaggregation and the more detailed IOp analysis with its results and some counterintuitive findings of the Nepalese reality that provide most of the value-added of this study for the empirical literature. The country seems to suffer from opportunity deprivation hindering individuals and HHs to perform to their fullest potential. Therefore, the Nepalese people have taken into hand their own destiny no matter their level of the socio-economic status. Geographical mobility, both domestic and international, are seen as an opportunity-equalising tool for the most disadvantaged, as well for the most advantaged. From the results, a rather dark picture is painted for Nepal, a country that is rich in cultures, natural resources, and biodiversity. As Bhattachan (2003) claims, "Nepal is a rich country inhabited by poor people". Today, the future of its people is blocked by inequality of opportunity across all social strata, yet, their opportunities should be endless.

In order to help Nepal leapfrog over this trap some policy suggestions are given. Behaviours and social discrimination will not change unless they stop severing as a socio-economic safety net and the yet fragile infant democratic Nepalese government can guarantee to serve as a social welfare state.

For this study to be more thorough and its policy implications to more effectively focus on the roots of the problems, some of this study's limitations and a broad band of possible research extensions were given.

Chapter 4 Inequality in Outcome and Inequality of Opportunity: Results Comparison and Literature Contributions

The goal of this chapter is to compare Chapter 2 on inequality of outcome and Chapter 3 on inequality of opportunity and highlight this study's contributions to the literature. This will be done first by looking at the various methodologies used in the inequality in outcome chapter. Second, the IOp methodology and its contributions will be stated. Third, this chapter highlights the general empirics and their literature contributions and, lastly, states the most important results of the previous two chapters.

Table 4.1 summarises all the key literature contributions of this study. The following sections will briefly discuss and explain what is given in the table. Please note that this table and this chapter only draw on the most important findings and contributions. The corresponding sections in this dissertation discuss, explain, and interpret these in more detail.

Table 4.1: Summary of Literature Contributions: Inequality in Outcome and Inequality of Opportunity

Literature	Contributions
4.1 Inequality in Outcome Methodology and Empirics	
Ample literature on income and consumption inequality (theory and empirics), but little consensus on how to aggregate economic outcome, which in turn influences the inequality indices' computation.	- Literature review on the measure of (economic) welfare, and justification for the utilisation of income and consumption as the most adequate economic welfare measures. - Literature and guideline review of the income and consumption expenditure aggregation methodologies - Adjusted aggregation methodology developed for the NLSS III dataset and the Nepal-specific context → important for inequality in outcome indices → important for subsequent IOp analysis - Thorough computation and numerous robustness checks (e.g. unit measurement checks, bootstrapped standard errors for indices, outlier checks, and adjusting for different costs of living) of the economic outcome aggregates, and aggregate results comparison. - Thorough computation and results comparison of various inequality in outcome indices. - Theoretical and empirical discussion on the use of the MLD as the reference inequality in outcome variable for the IOp analyses
4.2 Inequality of Opportunity Methodology	
IOp Methodology (three main strands) 1. Linear regression for economic outcome in terms of circumstances and efforts. Use of hypothetical distributions. Consequent earning decomposition. Direct and indirect effect of circumstances (e.g. Bourguignon et al., 2007) 2. Between-group inequality decomposition. Ex-ante and ex-post	**IOp analyses** 1. Combination of IOp analyses approaches 1 to 2 and inspiration from Ferreira and Gignoux (2011) and **adaptations to the NLSS III dataset** and **Nepal-specific country context** (e.g. choice of circumstances and categorical variables). Direct ex-ante approach of the compensation principle. 2. **Parametric and non-parametric** IOp indices 3. **Additional robustness checks** (e.g. regression tests and results comparison between different level analyses)

measures of IOp (e.g. Checci and Peragine, 2010) 3. Stochastic dominance comparison of distributions by type (e.g Lefranc et al., 2008)	4. **Population sample diversification and relaxation of data constraints** (e.g. not just HH heads and spouses but all HH members)
4.3 Inequality in Outcome and Inequality of Opportunity: General Empirics	
- Most literature **focuses on the developed world** e.g. the US and EU economies. Few analyses on the developing world due to data constraints. Even WB's HOI Report mainly focuses on LAC and sub-Saharan Africa. - **Caste societies**: main focus on India (by province). - **IOp: national indices** and then within inequality decomposition. - **Focus on specific opportunities** e.g. educational, health, or labour access opportunities.	1. **Geographic extension**: developing Asia, Nepal. Relatively exhaustive HH level dataset → Nepal-specific context. 2. **Caste society extension**. The only other IOp studies that include caste belonging are those on India. 3. **IOp level disaggregation**: National IOp analysis (as in other empirical literature) for Nepal, but further disaggregated computation by Development Region, urban vs. rural areas, and income and consumption groupings (quarters). In total: four levels of IOp analyses. Within-group and between-group analyses. 4. **Overall IOp indices** including various opportunity depriving/enhancing elements (i.e. circumstances). Further depth is added with RF-OLS regression analyses of circumstances on economic outcome and circumstance specific shares (partial IORs). Isolation of opportunity profiles by type and by economic advantage.
4.4 Inequality in Outcome and Inequality of Opportunity: Results	
<u>Analyses Disaggregation Results</u> 5. Important **divergences in terms of economic welfare across the country**. Variance in both level and concentration of economic welfare. 6. **Importance of geographical disaggregation** in understanding the local reality. "One size fits all policy" would be ineffective. 7. **IOp national level relatively low** (26-32%) compared to the literature → geographical disaggregation paints a different picture: important divergences by geographical region (e.g. urban (34-44%) vs. rural (17-22%)) and type of individuals. 8. **Consumption quarters:** bottom 25% of the population face a bare 3-9% vs. top population quarter 7-18% of within-group opportunity deprivation out of overall outcome inequality. **<u>Counterintuitive Results</u>** 6. **Gender**: mostly insignificant in IOp and OLS analyses, but important gender gap for the most impoverished in the central region. 7. **Parental education**: father's education tends to be more important than mother's education 8. **Religious minority**: Buddhists are afforded an elite and opportunity-enhancing status 9. **Caste group belonging**: still matters, yet more in urban than rural areas 10. **Rural birth region**: at the national level, being born in the countryside negatively affects economic outcome, but when disaggregated, the most disadvantaged types suffer less in rural areas. **<u>Other Interesting Results</u>** 4. In general, across all analyses, the mean economic outcome drops most severely when **four or more opportunity-depriving circumstances** are accumulated. Disproportionately negative in the central region and urban areas → should either emigrate or remain in the countryside. 5. Nepal is in a **poverty and development trap** due, in part, to opportunity deprivation → migration and remittance dependence, foreign aid dependence, social networks acting as socio-economic safety nets, etc. 6. **Geographical mobility as opportunity-equalising instrument**: both the most disadvantaged and the most advantaged migrate (domestically and internationally) → vicious cycle 7. **Sex ratio:** improves as the share of mean economic revenue increases. i.e. the higher the mean income level, the higher the gender balance → dependence of poorer households on migrational labour revenue.	

4.1. Inequality in Outcome Methodology and Literature Contributions

The study of inequality is at the core of economic development and the literature on inequality in economic outcome is ample both at the theoretical and empirical levels. Much of it is considered as acquis; however, it was to the author's surprise to uncover that there is little consensus on the methodology used to aggregate the variables that then serve as a key reference in calculating the inequality indicators. In other words, there is little consensus in the literature on how to aggregate, for instance, income and consumption expenditure. This study investigates a) the literature on the measure of (economic) welfare to justify the choice of income and consumption expenditure as economic welfare proxies. Most of the IOp literature does not debate as to why these two economic welfare measures are used in their analyses, while this study places significant emphasis on the arguments for and against them.

b) This study undertook a thorough literature review and comparison of the most prominent guidelines on the aggregation methodologies of income and consumption in order to optimise the approach for the NLSS III dataset and the Nepal-specific country context. None of the IOp studies the author came across discuss their aggregation methodologies for the economic welfare measures, despite them also using microlevel data and the important affect the aggregation methodology can have on the empirical analyses.

c) An adjusted aggregation methodology was proposed for both income and consumption. These approaches take into consideration i) the particularities of the NLSS III dataset, and ii) the Nepal country-specific context. Such elements include adjustments to regional living standards, the inclusion or exclusion of health and educational expenses, and unit measurement conversions.

d) Variations of the income and consumption aggregates were calculated following the recommendations in the literature and the author's adjusted approaches. These results were checked for robustness and consistency at every level of the aggregation process (e.g. unit measurement checks, bootstrapped standard errors for indices, outlier checks, and adjusting for different costs of living). The aggregate results vary significantly depending on the aggregation methodology used. This underpins the validity of the chosen fine-tuned methodologies for the author's recommended consumption and income aggregates.

e) It is important to choose the appropriate aggregation methodology and economic outcome measures carefully for the following inequality in outcome and IOp analyses. Previous IOp studies do not appear to place emphasis on these choices and simply take the aggregate levels of income and/or consumption for granted without discussing them. These have, however, significant impacts on the further inequality (in outcome and opportunity) analyses. The impacts were highlighted by empirically showing the effects that the aggregation methodologies have on i) the amplitude of the chosen inequality measures, ii) the divergence between the inequality measures for a given economic outcome variable, and iii) the divergence between the inequality indices, depending on whether consumption or income serve as the main economic welfare measures.

f) A thorough literature review of the most commonly adhered to and referenced inequality in outcome measures was done. The most suitable inequality in outcome indicator was isolated through both a theoretical and empirical investigation. The chosen inequality in outcome index is the Mean Logarithmic Deviation (MLD). This is a commonly used index in the inequality in outcome literature. It also serves as the inequality in outcome reference index for Ferreira and Gignoux (2011). However, they also do not explain nor greatly discuss the extent of the use of the MLD. This study contributes in giving a more elaborate debate on the trade-offs of using the MLD. This is important since the most intuitive and relevant IOp results (particularly the Inequality of Opportunity Ratio (IOR)) are obtained and interpreted by using the MLD as a reference. Two of the key advantages of the MLD are that it is the only measure of the Generalised Entropy (GE) inequality indices to satisfy the path independence axiom and that it is decomposable between income (consumption) sources and between individuals.

4.2. Inequality of Opportunity Methodology and Literature Contributions

This dissertation is primarily an empirical application of an IOp model developed and often referred to by FG (2011). The advantage of the model is that it combines the first two of the three main methodological strands of the empirical IOp literature. These are, first, a linear regression model for economic outcome in terms of circumstances and efforts, which means that IOp is both directly and indirectly captured. This complicated model makes use of hypothetical distributions and allows the decomposition of earnings (Bourguignon et al., 2007). Second, Checci and Peragine (2010) do a between-group inequality decomposition, which allows for both ex-ante and ex-post measures of IOp. Last, Lefranc et al. (2008) compare earning distributions by type by applying a stochastic dominance model.

So, this study combines the first two of these three main IOp approaches, inspired by the model developed by FG (2011). However, numerous adaptations are done in order optimise the operationalization of their model to the NLSS III dataset and the Nepalese country-specific context. For instance, the choice of circumstance variables varies. Religion, caste group belonging, and ecological belt stratification as the birth region are included. So, more circumstance variables are included, arguably capturing a broader field of exogenous environmental variables. Also, the circumstantial categorical variables are divided differently according to the Nepalese context and observation distribution. For instance, parental education is divided into only two categories rather than three in FG (2011): no education or primary education incomplete, and primary education completed and above.

Various additional robustness checks at all stages of the analyses are applied. These include, for instance, several regression tests and thorough results comparisons between the different levels of analyses (statistical significance tests). Also, both a parametric and non-parametric approach for the computation of the scalar IOp indices are used in order to check for additional robustness.

Furthermore, a few of the data constraints originally applied by FG (2011) are relaxed. While at the national-level analyses, one population sample has been constrained using the same conditions as applied by FG (2011) in order to allow for results comparability, two other population samples are analysed. Both extend the inclusion of observations to members of the HH other than the HH head and spouse. This increases the sizes of the population sample and the study's viability.

4.3. Inequality in Outcome and Inequality of Opportunity: General Empirics and Literature Contributions

Most of the IOp empirical literature focuses on developed countries, such as the US or European economies, as can be seen in Table 5.3 in the appendix, which summarises the main empirical IOp literature (e.g. Chetty et al. 2018; Checchi and Peragine, 2010; Brunori, 2016; Lefranc et al. 2008). Relatively little empirical literature exists for developing countries. There are a number of studies on Latin American countries (e.g. Ferreira and Menendez, 2007; FG, 2011;), few on African and Middle Eastern countries (e.g. Barros et al., 2009; Hassine and Zeufack, 2015; Hassine, 2011, Dabalen et al., 2015; Ferreira et al., 2011), and virtually none on Asia (e.g. United National ESCAP 2018; Asadullah and Yalonetzky, 2012; Singh, 2010). Even the World Bank's Human Opportunity report mainly focuses on Latin American countries and sub-Saharan Africa (e.g. Molinas et al., 2010). The scarcity of empirical IOp research is mainly due to the lack of availability of exhaustive and rich microlevel datasets. Most of them particularly lack information on parental background. This study contributes to the literature by being an empirical extension to developing Asia, namely Nepal. Access to a relatively rich micro dataset has allowed the author to exploit it for an in-depth analysis of IOp.

The benefit of this geographical empirical extension is twofold. First, it is an additional Asian country that is analysed. Second, it is a caste society. The only other IOp studies that explicitly include caste group belonging focus on India (e.g. Asadullah and Yalonetzky, 2012; Rama, 2016; Singh, 2010). Yet there are many other caste, or caste-like, societies such as Sri-Lanka, Pakistan, and Bangladesh.

The IOp literature tends to compute nation-wide inequality in opportunity profiles for types of people, which are then disaggregated in terms of their circumstances or earnings or by geographical factors. This study, however, adds three layers of analyses to the national level IOp investigation. This is done by also studying the amplitude of IOp at two different geographical levels (the Development Region level and the urban versus rural area level) and at a population groupings level (division of the population into income and consumption quarters). In other words, there are four levels of analyses for the study of IOp in Nepal that complement each other and add depth to the study. At all levels, within and between group IOp is tested. This disaggregation in this study is Nepal-specific and cannot directly be compared to other countries or regions. Some general suggestions may, however, be derived. More on this in the results' value-added section.

Some IOp literature focuses on specific sub-categories of opportunity deprivation, such as inequality in educational or health opportunity (e.g. Assaad et al. 2014; Ferreira and Gignoux, 2014; Asadullah and Yalonetzky, 2012). This study takes another empirical approach due to the richness of the dataset in computing overall IOp indices that try to quantify the globality of opportunity deprivation in relation to total inequality of outcome. In other words, it tries to quantify how much of the total economic welfare inequality is due to opportunity deprivation. Then, *opportunity deprivation profiles* are analysed and an in-depth investigation in the importance of each circumstance variable is done by using RF-OLS regressions and calculating *partial IORs*. The emphasis here is very different to what the sub-categorical opportunity studies try to show.

4.4. Inequality in Outcome and Inequality of Opportunity: Results Comparison and Literature Contributions

The goal of this section is to highlight the globality of the results in this dissertation. It combines the most striking findings from the inequality in outcome chapter and the IOp chapter. While both chapters state, discuss, and interpret their results, respectively, this chapter only draws on the most important and the counterintuitive results obtained. To do so, first some remarks are made on the national-level results and the importance of disaggregating them into three additional levels of analyses. Second, the main five counterintuitive findings are discussed. Last, a few general comments are made about the globality of the results and some of the implications that they induce.

1) Analyses disaggregation results

First, there are **important divergences in terms of economic welfare** across the country, these divergences are in both the concentration and level of economic welfare. At the national level and in terms of consumption expenditure, over 4% of the population i.e. 1.14 million people, live below the 1USD/day poverty line. Over 85% of Nepalese live on below 2USD/day. Disaggregating this by Development Region, "only" 55% of people in the central region live below the 2USD/day poverty line, while this reaches more than 98% of people living in the far-western and mid-western regions.

In general, of the top earning and consuming quarters of the population, more than 48% and 54%, respectively, live in the central region. The top quarter earners have almost 70% of the national income share, while the top consumers about 56%. For the bottom quarters, both earners and consumers hold less than 3% and 9% of the national total, respectively.

Moreover, the **geographical disaggregation** of the IOp analyses highlights the importance in understanding the local reality. It becomes evident that a "one size fits all" policy would be ineffective. The level of **IOp at the national level** is relatively low compared to in the literature (26% to 32%). However, the geographical disaggregation paints a very different picture. For instance, at the urban versus rural area level analysis, in urban areas, more than 34% to 44% of overall inequality of economic welfare is due to unequal opportunities. In rural areas, this is only about 17% to 22%. At the Development Region level, especially the central region faces by far the highest level of IOp (37% to 44%). The results suggest that in particular, men tend to migrate to the urban areas and central region hoping to seize opportunities and making a living. Yet, that is where they suffer the most from opportunity deprivation. The urban rural level IOp analysis suggests a disproportionate "urban circumstantial amplification" for negatively impacting circumstances. In the countryside "everyone is poor".

Furthermore, it is evident that **the more circumstances one accumulates** that negatively impact on opportunities, the more an individual is discriminated against and the lower the expected mean per capita income and consumption. In general, for individuals who accumulate four or more discriminatory circumstances, there appears to be no difference in gender or having been born in the mountain, hill, or Terai areas. When referring to the partial IOR estimates, at the national level and for the consumption panel, ethnic and religious minority belonging, parental education, and being born in a rural area account for up to 76% of overall outcome inequality suggesting their importance for accessing opportunities. The national-level partial IORs are relatively low and homogenous compared to those when they are geographically disaggregated. For instance, the same categorical circumstance variables explain more than 107% of inequality in the central region, while barely 30% in the eastern region.

When looking at the Nepalese population divided into quarters by consumption levels, it can be seen that the bottom quarter faces only 3% to 9% compared to a 7% to 18% degree of within-group opportunity deprivation out of overall consumption inequality. When this is combined with the urban IOR estimates (urban IOR is 34% to 44%, rural IOR is 17% to 22%) the results suggest that for the **limited opportunities available in Nepal**, the most privileged in society are in fierce competition and need to maximise the mobilisation of social capital. Social networks do not only act as a means of socio-economic safety nets (i.e. insurance), but are also "an essential element for youth and adults who seek employment and the quality of these networks are often linked to levels of income and social status" (Ezemenari and Joshi, 2019, p. 6).

The people of the **bottom income quarter** appear to suffer greatly from opportunity deprivation (2% to 54% out of total income inequality). This result needs to be taken cautiously due to the large within-group variance of income levels (yet checked for outliers, etc. during the aggregation process). However, it does suggest that some of the poorest of the poor are in major debt and are trapped in a downward spiral. During the field trip, various possible explanations were unveiled for this. One factor that appears to be common is that particularly the most disadvantaged need to borrow capital if they wish to emigrate in the hope of improving their economic welfare. Emigration, however, does neither guarantee higher earnings abroad nor guarantee the family back home will receive any remittances, and if they do, they are received at irregular intervals. This traps the remaining family members in a downward spiral where they cannot pay back their interests. So, they continue to borrow more, etc. and this further aggravates their already stark opportunity deprivation. The Ministry of Labour and Employment report (MLE, 2018) highlights that despite the overall volume of remittances being large (27% of the fiscal year 2016/17 GDP), the individual level is as low as 80,462 NRs per annum (about 1,000 USD). More than 92% of Nepalese emigrants use recruiting agencies to apply for labour permits to allow them to emigrate. The report points out that these are only the official numbers and that the rates the agencies charge are often horrendous (MLE, 2018). From personal exchanges with women at KOPILA Nepal (NGO for the protection of human rights and psychosocial well-being; various informal interviews, October 2018) the author was told that some are charged

400USD or more (equivalent to their annual income) for the assistance of an agency. This is to conclude that while the bottom income quarter IOp results need to be looked at with caution, they do back up the fact that the most deprived need to indebt themselves to emigrate and are often caught in a vicious downward spiral.

2) Counterintuitive results and brief interpretations

There are five main counterintuitive results from this study's analyses. These do vary in amplitude across and within the different levels of analyses, so here only the general trends are commented on.

First, the **gender** circumstance variables are **low or completely statistically insignificant** across all levels of analyses. This is despite international organisations, local Nepalese non-governmental organisations, and the Nepalese government stressing the need to fight for gender equality in Nepal (e.g. Bennett, 2005; KOPILA Nepal, 2018; GIZ, 2013; Election Commission, 2013). There are several explanations for this. One is the partial endogeneity and partial exogeneity of gender in the NLSS dataset. Furthermore, the sex ratio of the analysed population samples are exceptionally low, overestimating the importance of men. However, at the national level, the gender RF-OLS coefficient is statistically significant, and being a woman explains over 15% of consumption inequality according to the partial IOR estimates suggesting its importance as an explanatory factor for opportunity access. Also, a clear gender gap exists amongst the most opportunity deprived in the mountain belt in the central region. The poorest of the most deprived in the centre are women who are born in the mountain belt, while the least poor of the most disadvantaged in the centre, are men in the mountain belt. Their mean level of consumption is more than 250% that of women.

Second, overall **father's education appears to be more important** than mother's education in determining the economic welfare of children. This is valid at the national level in rural areas and in four out of five Development Regions (apart from the mid-west). There are some possible explanations for this. First, the combination of male outmigration and the practice of joint families means that the remaining mothers are overburdened in trying to look after the entire family, the in-laws, the children, and then also further overburdened in trying to generate an economic livelihood (uncertainty of remittances and the regularity of them; , 2018). This means that often it is the remaining other male family members, such as grandfathers or brothers, who take many of the daily life decisions. Second, modern digital communication means that de facto, fathers are still the main decision makers, especially in regards to important resource allocation, such as higher education expenses. Third and partially surprising, the father's education level is more important in the central region (which is most developed and hosts the capital). Yet, this is also where most economic opportunities lie, the sex ratio is the highest, and the HH sizes are the largest, thus indicating a higher share of the presence of a male dominant figure who takes decisions.

The RF-OLS regression results, however, indicate that in urban centres, the mother's education is more important than that the father's (the difference between them is statistically significant). Also, the partial IOR results indicate that the mother's education explains more than 25% of overall inequality versus 21% for the father's in urban areas, suggesting the importance of parental education in guaranteeing access to opportunities for their children. As expected, the exact opposite is the case for rural areas. Furthermore, the access to education deteriorates disproportionately more for fathers than mothers the further away a rural area is from the capital.

The mothers' education levels matter most in urban areas partially because "single" mothers move to other family members in the city for their children's education, accepting the fact that discrimination rates are disproportionately high in urban centres. This is particularly the case for the

most disadvantaged families, who thus hope that the enhanced education level of the child will lead to their better future for them. The World Bank (2016) confirms this and emphasises the trend of urban migration in order to access certain amenities, such as education and health care.

The large geographical divergence in the prevalence and amplitude of parental education can be seen when referring to the partial IOR estimates. At the national level, parental education accounts for 31% of consumption inequality (similar to the far-west region, for example), while in urban areas it is over 45% and in the eastern region less than 8%. This suggests the importance of parental education in guaranteeing the accessibility to opportunities.

Third, despite **Buddhist** being a religious minority, they are afforded an elite and opportunity-enhancing status. In Buddhism, all living things are equal, and this largely allows people to "break free" from their inborn caste belonging. Also, being a Buddhist monk grants access to discipline and better education than what is available publicly. Yet, conversion to Buddhism has slowed down as Nepalese society is developing and behaviours change to no longer discriminate based on caste grouping or ethnicity.

Despite Nepal having one of the most progressive human rights and equality enhancing legislative frameworks and having officially abolished the caste system, **caste belonging still matters**. Unexpectedly, this is all the more the case in urban than rural areas. At the national level, ethnic group belonging accounts for "only" 14% of inequality, while for urban areas it is over 34% and in the central region 23%. The 34% partial IOR estimate for urban areas is the highest partial IOR across all levels of analyses and suggests a continued residential and physical segregation of ethnic minorities and caste groups in urban centres. There are several explanations for the continued importance of caste and ethnic group belonging across the country in explaining opportunities and inequalities in economic welfare. First, the data constraints applied (30 to 49 year olds) mean that the analysed population is of a slightly older, possibly more conservative generation. This is aggravated by the fact that the NLSS III dataset is from 2011, meaning that the analysed individuals were born between 1962 and 1981 in an even less developed Nepal. Second, as mentioned before, particularly mother's in the urban centres (higher rate of joint families) are overburdened, and the children are brought up by the older (and often male) more conservative HH members (e.g. grandparents or bigger brothers). Third, the families that are usually divided by emigration tend to be the poorer and less privileged ones, in other words the more traditional ones with relatively lower education levels. Last, the lower the economic welfare of a family the more dependent they are on their within-group network to act as a socio-economic safety net. So, automatically, anyone from within the group is preferred to someone outside the group when it comes to any socio-economic decisions. Ezemenari and Joshi (2019, p. 6) state that "personal networks are an essential element for youth and adults who seek employment and the quality of these networks are often linked to levels of income and social status".

Third, at the national level, **being born in a rural area** negatively impacts on economic outcome. When disaggregated, however, people of the most opportunity-deprived types suffer less when they are born in rural areas e.g. born in the Terai region (positive impact, urban-rural analysis) or mid-west and far-west regions (much lower negative impact than in the central region).

3) General remarks on the globality of the results and their possible implications

First, in general, across all analyses, mean economic outcome drops most severely when **four or more opportunity depriving circumstances** are accumulated. The effect of negatively impacting circumstances is disproportionately amplified in the central region and urban areas. These results

confirm those by the World Bank (2016, p.13) that "the pull of a dynamic urban economy generating good jobs appears to be largely non-existent." This would suggest that if one belongs to the most disadvantaged type, one should a) either emigrate, or b) move to / remain in the countryside. It also suggests that the government should seize the chance of its current federalisation process to decentralise its economic structures and more vigorously try to invest in rural areas in order to promote the socio-economic and inclusive development of the countryside.

Second, **Nepal seems to be trapped**, trapped in poverty and trapped at a stage of development where all people independent of socio-economic status greatly suffer from opportunity deprivation. The results of this study suggest that it is particularly the high levels of opportunity deprivation that hinders individuals and HHs to perform to their fullest potential, which in turn slows down the country's development. Largely due to the failure of the government in providing opportunities and access to them (e.g. economic, welfare, social), the Nepalese people have taken their destinies into their own hands. Geographical mobility, both domestic and international, is seen as an opportunity-equalising tool for the most disadvantaged, as well for the most advantaged. This has encouraged a multi-level downward spirals for development, all of which are interconnected. The more time passes the harder it will be to reverse their effects. The first spiral consists of the most disadvantaged emigrating as low-skilled labourers, often indebting themselves and overburdening the left mothers due to the uncertainty of receiving remittances (MLE, 2018). Furthermore, the poorest cannot afford to send the entire family abroad. Thus, the impacts on the livelihoods of children for being brought up by a single overburdened parent are yet to be studied. Also, this means that the most conservative and poorly educated remain in Nepal.

The second spiral is driven by most educated and skilled who emigrate, leading to not just brain drain (MLE, 2018), but also the loss of well-connected social networks that could otherwise facilitate the development of a dynamic and diverse economic system. Furthermore, emigration has an element of prestige attached to it for the advantaged. Being able to afford to study abroad (even if it is through a scholarship program), working in a regular office job wearing a suit, or maybe being able to buy a motorbike or even a car, comes with great prestige at home and recognition.

The third downward spiral consists of the long-term socio-economic consequences of the development of the dual education system (public and private). Suvedi et al. (2014), for instance, elaborate in length on the changes in Nepal's education system and the impacts of its privatisation. A large share of remittances are invested into children's private education. This decreases the responsibility of public institutions to provide for high quality and accessible education at the schooling level but also especially at the higher education level (Ghimire, 2018). This further incentivises foreign aid to flow into the public schooling sector, in return further aggravating the situation.

The last spiral consists of Nepal having integrated remittances as a key pillar of its development strategy. It has done so, for example, by facilitating the possibility to obtain emigration labour permits, aggressively elaborating bilateral labour agreements (e.g. fair treatment of labourers abroad), or easing the flow of small sum financial means back into the country (MLE, 2018). Despite this of course improving the welfare and living conditions of Nepalese labourers abroad, it further incentivises more people to go abroad. Furthermore, the World Bank (2016) argues that Nepal's dependence on remittances has two major adverse effects on its economic. First, it has caused the appreciation of the real exchange rate, consequently decreasing the competitiveness of the infant manufacturing sectors and further undermining job creation. Second, they argue that the continued outflow of young Nepalese could have disincentivised policymakers to push for urgently needed mid and long-term reforms. Yet, Ezemenari and Joshi (2019, Part B) discuss a whole range of challenges but also opportunities created by remittances. On the one hand, for instance, remittances will continue to drive private investment and consumption as a substantial part of GDP growth, and build up foreign exchange reserves. On the

other hand, for instance, a sudden reduction in remittances would significantly lower deposits and thus credit availability for the economy.

All of these four downward spirals are inter-connected and further disincentivise the creation of a diversified and dynamic domestic economy. Key will be to guarantee political stability and will to renounce some of the immediate short-term interests for a long-term and solid-based socio-economic system.

Third, the decade-long failure of the political institutions to deliver increased social and economic mobility has forced people to take action themselves. Consequently, everyone but especially the most disadvantaged and the most advantaged groups of people use **geographical mobility**, both domestic and international, as an opportunity-equalising instrument.

Foremost, especially the **most opportunity deprived** are forced to migrate. A) They migrate domestically to urban centres for education purposes. Yet, in urban centres opportunity deprivation rates are the highest in the country, and this is disproportionately the case for the most discriminated. So, once the most disadvantaged receive some basic education, they try to emigrate. B) The most disadvantaged migrate internationally as low-skilled labourers (20-25% of the population; MLE, 2018) with the hopes to then be able to send back remittances. In turn, remittances have become a major share of Nepal's GDP (around 25%; World Bank, 2011), making them a substantial pillar to Nepal's economic development strategy. It can be argued that a secondary positive effect of these low-skilled labour emigrants is a brain gain upon their return. However, Nepal has entered a vicious circle economic development spiral with too much emphasis on remittances and international aid, so that when the "newly trained" medium-to-high-skill emigrants return, they cannot find employment in Nepal. As previously mentioned, the World Bank (2016) highlight that Nepal's dependence on remittances has two major adverse effects on its economy. First, the harmful effects of an appreciated real exchange rate on international competitiveness and job creation and, second, the disincentive for policymakers to implement mid- and long-term structural economic reforms.

Additionally, the **more advantaged groups** also migrate. A) They migrate domestically, and especially to the urban centres and the capital to seek higher education. But, the employability of young Nepalese and a serious lack of job opportunities in Nepal make it very difficult for them to make a living. This study's results indicate that opportunity deprivation rates are the highest in the urban centres. Arguably, Nepal's limited economic development and having a centralised state means that there are more, albeit still few, opportunities in urban areas and especially the capital. However, this also means, that it is easier to discriminate on circumstances and the greater number of "opportunity privileged" individuals in urban centres are in fierce competition and must use all their social capital to seize opportunities. In other words, the high population density in urban centres for relatively few economic opportunities means that these are reserved for only the most opportunity-privileged people. Individuals who accumulate opportunity depriving circumstances are left with disproportionately less opportunities. So, they also have to emigrate eventually (or move back to the countryside). B) The most advantaged groups also emigrate internationally, 1) to seek high-quality university education, 2) to seize quality employment, 3) to escape the conservative mind-set and societal pressure at home e.g. for out of caste system marriages, and/or 4) prestige.

To sum up, due to opportunity deprivation and the inability of the Nepalese government to tackle these, all people independent of socio-economic status see use geographical mobility as an opportunity-equalising instrument.

Last, the **sex ratio improves** as the share of the mean economic outcome increases. In other words, the higher the level of economic welfare, the higher the gender balance. This can be a) because the richer HHs are more educated (positive correlation between mother's and father's education) and thus there being higher gender equality, and b) because particularly the poorer HHs depend on migrational labour revenue.

To conclude, this chapter has shown that this dissertation contributes to the literature in four main ways. First, by digging into the inequality of economic outcome literature and drawing up an adjusted aggregation methodology for the economic outcome variables and discussing the most appropriate measure of inequality in outcome for the subsequent IOp analyses. Second, this chapter highlighted the contributions to the IOp literature in terms of methodology and especially on how to adapt the FG (2011) direct ex-ante approach to studying IOp to the NLSS dataset and the Nepal-specific country context. Third, the contributions to the literature on inequality in outcome and the IOp analyses at the general empirical level are given. Last, the most striking results from both the inequality of outcome and the IOp analyses are combined and highlighted.

Chapter 5 Conclusion

In order to fully analyse Nepal for the existence of IOp, this dissertation consists of two main analytical chapters. Chapter 2 investigated the level of inequality in outcome (income and consumption expenditure), which served as the basis for the subsequent analysis of IOp in Nepal in Chapter 3. Chapter 4 compared and highlighted the literature contributions of the previous two chapters. Chapter 5 concludes this dissertation.

The introduction chapter outlined the research questions, hypotheses, literature contributions, and findings for both the inequality in outcome and the IOp analyses.

Chapter 2 showed that the level of inequality in economic welfare can be analysed in many ways that are complementary to each other. First, inequality can be analysed in a descriptive manner by looking at the distribution of income across a society by population and by geographical dispersion, for instance. Second, inequality can be analysed using more sophisticated statistical tools, drawing on the distributional disparity of economic welfare using, for instance, the Gini or Theil indices. They provide additional information using the distributional characteristics for understanding the disparity in economic welfare across a society. These two steps consider that the underlying concept of inequality is the uneven or "unfair" distribution of outcome across a society using, for example, income as a monetary measure for economic welfare. However, one can argue that not only income and consumption, but also other elements such as life expectancy, health status, or education level, contribute to one's welfare or living standard. Just like the concept of welfare, the concept of inequality is also multifaceted, and one must move beyond a unilateral monetary measure to a more complete measure and analysis of it. So, while the analysis of economic disparity is crucial in serving as a starting point for any inequality analysis, other inequalities must be considered. These include, for instance, educational and health inequality, nutritional inequality, social mobility inequality, or various market access inequalities (e.g. labour or credit market).

The purpose of Chapter 2 was threefold. First, to compute the most suitable economic welfare variables using the most suitable aggregation methodology. Consumption expenditure and income were identified as the most commonly used economic welfare proxies, both bearing their advantages and disadvantages. Of the two, consumption expenditure is the preferred measure given the Nepalese context, because it is seen to more adequately represent the current living standards of the people than income. Also, income bears a temporal component (current and future welfare) and therefore is arguably less appropriate for this dissertation's subsequent IOp analyses. It was highlighted that great care must be taken when aggregating the economic welfare variables. The author computed consumption and income aggregates following various guidelines and concluded which elements in the aggregation process of them are the most suitable given the Nepalese context and the nature of the NLSS III dataset. As a result, one "ideal" consumption and one "ideal" income aggregate were proposed using "optimised" aggregation methodologies. It is crucial to understand how the economic outcome variables were computed when interpreting the consequent descriptive distribution disparity and the further statistical econometric results for both the inequality in outcome indices and the IOp analysis. In other words, different aggregation methodologies have different impacts on the statistical analysis and their results.

The second purpose of Chapter 2 was to describe the distribution and disparity of economic outcome across Nepal using both graphical illustrations and descriptive statistics. To further complement the nation-wide disparity description, distributional inequalities were broken down by population quintiles and by geographic zones. A large degree of disparity was noted for all outcome variables at the national level and at the geographical breakdown level. However, the degree of

disparity depends, to some extent, also on the aggregation methodology used to compute the outcome variable.

The third purpose of Chapter 2 was to identify the most universally recognised inequality in outcome indices, weigh their pros and cons, and to estimate and describe them. The most commonly used Generalised Entropy and Atkinson Inequality Indices were computed. Different inequality in outcome indices place emphasis on different parts of the distribution, thus giving different results. Judging from the literature review and the inequality in outcome analysis, the Mean Logarithmic Deviation (one of the Generalised Entropy Indices) was identified as the most suitable inequality in outcome reference variable for the subsequent IOp analyses. Two of its key advantages are that it is the only measure of the GE indices to satisfy the path independence axiom and that it is decomposable between income sources and between individuals.

To conclude, Chapter 2 mainly described the distribution, disparity, and inequality of income and consumption in Nepal. However, it did not manage to explain these with rigour. It is difficult to generalise about the fairness and social justice of a society using purely descriptive disparity and single digit inequality in outcome measures. Inequality is a multifaceted and complex concept and in order to explain it, more complete measures of inequality need to be applied. Chapter 3 proposes a measure of inequality that seeks to complement the inequality in outcome analysis. It aims at explaining and quantifying the factors that most impact on the inequality in outcome. To do so, a measure for IOp is proposed, which first quantifies the proportion of inequality in outcome that is due to opportunity deprivation, and second decomposes this IOp in terms of explanatory factors.

Chapter 3 tries to build on the inequality literature, and more precisely on a relatively recent sub-branch of it: IOp. Inequality in opportunities is the core thought behind development economics. Most topics in development economics refer to the key concept of opportunities and access to them, such as credit market access, the inequality-poverty-economic growth triangle, or child labour. Yet, the term *inequality of opportunity* only became a reference with Roemer's 1998 benchmark book establishing the field of opportunity economics. Other concepts such as intergenerational (e.g. income or educational) mobility, meritocracy or just inequality could be perceived as closely related, but distinct from the concept of IOp. "Opportunity economists" (e.g. Bourguignon et al., 2007; Checci and Peragine, 2010) argue that these terminologies are partially incoherent because they include both static and dynamic elements as well as inequalities in outcome and opportunity that arise due to individual choice and exogenous factors.

The literature on inequality in outcome (e.g. income) is abundant. It may, however, be argued that the distribution of particular outcomes is not appropriate for assessing the fairness of a social system or an allocation (e.g. Sen 1985, Arneson 1989). Inequality that arises due to differences in choice (attributes for which individuals can be held responsible and accountable) and therefore different outcomes is more socially and ethically acceptable. In other words, income gaps that arise from differences in *efforts* are generally less objectionable than those that are due to differences in race, parental background, or gender, for example. The goal of the decomposition of IOp is to single out to what extent inequality is due to unequal *opportunities*.

The importance of the study of IOp for the public, policy makers, and economists was stressed by the shift in public opinions about inequalities in outcome. Also, a gradual shift in the literature of distributive justice away from the traditional analysis of inequalities in outcome to inequalities in economic inputs was highlighted. However, the academic literature on IOp is relatively recent and scarce on a theoretical level and even more so on an empirical level. Most of the scarce empirical literature focuses on developed countries, such as the US and some European economies. This is because there are numerous challenges and trade-offs with the IOp concept initially developed by Roemer and then further extended by others, one main difficulty being the challenge to operationalise the IOp theory, particularly through the difficulty in measuring efforts.

Yet, Chapter 3 tries to complement the inequality in outcome analysis with an in-depth analysis of IOp in Nepal. This study is mainly an empirical study aimed at understanding the current IOp situation in Nepal rather than being a theoretical investigation. To do so, this study replicates the direct ex-ante model of the compensation principle developed by FG (2011). Some methodological and variable adaptations of the model are done in order to suit the NLSS III dataset and fit into the Nepalese socio-economic, political, and cultural context. This study, also, excludes the heavily debated effort variable from the IOp analyses. In fact, much of the empirical literature excludes efforts because the inclusion of it requires the necessary independence assumption between circumstances and efforts, which is arguably seen as too strong. Additionally, since the IOp indices computed in this dissertation arguably satisfy the path independence axiom, the ex-ante approach (which does not require the observance and measurement of the effort variable) delivers reliable and identical results to the ex-post approach (which includes effort).

The Nepalese NLSS III dataset was tested for IOp, and the opportunity deprivation profiles of individuals were singled out and cross-compared at four different analytical levels. The first is the national level analysis, which was then disaggregated into Nepal's Development Regions. Further analyses disaggregated the national level into urban and rural areas. Last, the opportunity deprivation of the population, which was divided into four economic welfare groups, was analysed.

Despite the national-level results being somewhat low compared to the ones of other developing countries in the literature, the results are alarming. However, particularly the geographical disaggregation indicates that there are striking divergences across the country, both in terms of opportunity deprivation and in terms of the types of people, that is the most opportunity deprived and the most advantaged. The disaggregation and the more detailed IOp analysis with its results and some counterintuitive findings of the Nepalese reality provide most of the value-added of this study to the empirical literature. Table 3.93 summarises the literature contributions of the inequality in outcome and the IOp chapters.

The results suggest that the level of IOp at the national level is lower than most in the literature, with opportunity deprivation accounting for 26% to 32% of overall inequality in outcome. The geographically disaggregated results, however, indicate a large divergence across the country in terms of amplitude and concentration. Urban areas, for instance, face opportunity deprivation levels reaching between 34% and 44% out of the total outcome inequality (vs. 17% to 22% in rural areas). Interestingly, the circumstantial decomposition suggests that negatively impacting factors (e.g. lower caste group belonging and the absence of parental education) have, first, a more ample effect in urban centres and this is, second, disproportionately aggravated by the accumulation of such negative factors. Moreover, the top 25% of the population by consumption faces a higher level of within-group IOp than the poorest 25% of the country.

The circumstance-specific analyses suggest that geographical disaggregation is crucial. There are some counterintuitive results. 1) Gender is consistently insignificant across all analyses. 2) Father's education seems to be more important on impacting on outcome than mother's education. 3) Buddhists, despite being a religious minority, are afforded an elite and opportunity enhancing status. 4) Caste group belonging still largely predicts the access to opportunities and more so in the urban than rural areas, despite Nepal having abandoned this societal stratification system decades ago. 5) At the national level, being born in a rural area appears to negatively impact on economic outcome. However, when disaggregated, the disadvantaged suffer less in rural areas and thus the rural birth region is seen as positive. In general, the most opportunity-deprived types are individuals who accumulate four or more of the opportunity-depriving circumstances, but the most disadvantaged types suffer disproportionately more in the central region and in urban areas. This suggests that either they should stay in the countryside or emigrate.

Furthermore, the country seems to be trapped, and that in part due to opportunity deprivation, where at all levels of the social stratification of the country, the Nepalese people have taken into hand their own destinies. Geographical mobility, both domestic and international, are seen as an

opportunity-equalising tool for the most disadvantaged, as well for the most advantaged. From the results, a rather dark picture is painted for Nepal, a country that is rich in cultures, natural resources, and biodiversity. As Bhattachan (2003) claims, "Nepal is a rich country inhabited by poor people". Today, the future of its people is blocked by inequality of opportunity across all social stratus, yet their opportunities should be endless.

In order to help Nepal leapfrog over this trap, some policy suggestions are given. These, to some extent cross-country comparable indices as well as their disaggregation allow for more precise and custom-tailored policies and projects to combat inequalities in opportunities and guarantee inclusive development. Consequently, these should enhance living standards, especially for the most impoverished and discriminated. Opportunity-equalising policies should focus on and encourage all individuals to engage in economic activities to the best of their abilities, in return boosting overall economic performance, social inclusion, and political stability, and enabling personal fulfillment. Yet, behaviours and social discrimination will not change unless they stop serving as a socio-economic safety net and the yet-fragile infant democratic Nepalese government can guarantee to serve as a social welfare state.

For this study to be more thorough and its policy implications to more effectively focus on the roots of the problems, some of this study's limitations and a broad band of possible research extensions are given.

Table 5.1: Summary of Literature Contributions: Inequality in Outcome and Inequality of Opportunity

Literature	Contributions
Inequality in Outcome Methodology and Empirics	
Ample literature on income and consumption inequality (theory and empirics), but little consensus on how to aggregate economic outcome, which in turn influences the inequality indices' computation.	- Literature review on the measure of (economic) welfare, and justification for the utilisation of income and consumption as the most adequate economic welfare measures. - Literature and guideline review of the income and consumption expenditure aggregation methodologies - Adjusted aggregation methodology developed for the NLSS III dataset and the Nepal-specific context → important for inequality in outcome indices → important for subsequent IOp analysis - Thorough computation and numerous robustness checks (e.g. unit measurement checks, bootstrapped standard errors for indices, outlier checks, and adjusting for different costs of living) of the economic outcome aggregates, and aggregate results comparison. - Thorough computation and results comparison of various inequality in outcome indices. - Theoretical and empirical discussion on the use of the MLD as the reference inequality in outcome variable for the IOp analyses
Inequality of Opportunity Methodology	

IOp Methodology (three main strands)	IOp analyses
1. Linear regression for economic outcome in terms of circumstances and efforts. Use of hypothetical distributions. Consequent earning decomposition. Direct and indirect effect of circumstances (e.g. Bourguignon et al., 2007) 2. Between-group inequality decomposition. Ex-ante and ex-post measures of IOp (e.g. Checci and Peragine, 2010) 3. Stochastic dominance comparison of distributions by type (e.g. Lefranc et al., 2008)	1. Combination of IOp analyses approaches 1 to 2 and inspiration from Ferreira and Gignoux (2011) and **adaptations to the NLSS III dataset** and **Nepal-specific country context** (e.g. choice of circumstances and categorical variables). Direct ex-ante approach of the compensation principle. 2. **Parametric and non-parametric** IOp indices 3. **Additional robustness checks** (e.g. regression tests and results comparison between different level analyses) 4. **Population sample diversification and relaxation of data constraints** (e.g. not just HH heads and spouses but all HH members)
Inequality in Outcome and Inequality of Opportunity: General Empirics	
- Most literature **focuses on the developed world** e.g. the US and EU economies. Few analyses on the developing world due to data constraints. Even WB's HOI Report mainly focuses on LAC and sub-Saharan Africa. - **Caste societies**: main focus on India (by province). - **IOp: national indices** and then within inequality decomposition. - **Focus on specific opportunities** e.g. educational, health, or labour access opportunities.	1. **Geographic extension**: developing Asia, Nepal. Relatively exhaustive HH level dataset → Nepal-specific context. 2. **Caste society extension**. The only other IOp studies that include caste belonging are those on India. 3. **IOp level disaggregation**: National IOp analysis (as in other empirical literature) for Nepal, but further disaggregated computation by Development Region, urban vs. rural areas, and income and consumption groupings (quarters). In total: four levels of IOp analyses. Within-group and between-group analyses. 4. **Overall IOp indices** including various opportunity depriving/enhancing elements (i.e. circumstances). Further depth is added with RF-OLS regression analyses of circumstances on economic outcome and circumstance specific shares (partial IORs). Isolation of opportunity profiles by type and by economic advantage.
Inequality in Outcome and Inequality of Opportunity: Results	
<u>Analyses Disaggregation Results</u> 1. Important **divergences in terms of economic welfare across the country**. Variance in both level and concentration of economic welfare. 2. **Importance of geographical disaggregation** in understanding the local reality. "One size fits all policy" would be ineffective. 3. **IOp national level relatively low** (26-32%) compared to the literature → geographical disaggregation paints a different picture: important divergences by geographical region (e.g. urban (34-44%) vs. rural (17-22%)) and type of individuals. 4. **Consumption quarters:** bottom 25% of the population face a bare 3-9% vs. top population quarter 7-18% of within-group opportunity deprivation out of overall outcome inequality. **<u>Counterintuitive Results</u>** 1. **Gender**: mostly insignificant in IOp and OLS analyses, but important gender gap for the most impoverished in the central region. 2. **Parental education**: father's education tends to be more important than mother's education 3. **Religious minority**: Buddhists are afforded an elite and opportunity-enhancing status 4. **Caste group belonging**: still matters, yet more in urban than rural areas 5. **Rural birth region**: at the national level, being born in the countryside negatively affects economic outcome, but when disaggregated, the most disadvantaged types suffer less in rural areas.	

Other Interesting Results

1. In general, across all analyses, the mean economic outcome drops most severely when **four or more opportunity-depriving circumstances** are accumulated. Disproportionately negative in the central region and urban areas → should either emigrate or remain in the countryside.
2. Nepal is in a **poverty and development trap** due, in part, to opportunity deprivation → migration and remittance dependence, foreign aid dependence, social networks acting as socio-economic safety nets, etc.
3. **Geographical mobility as opportunity-equalising instrument**: both the most disadvantaged and the most advantaged migrate (domestically and internationally) → vicious cycle
4. **Sex ratio:** improves as the share of mean economic revenue increases. i.e. the higher the mean income level, the higher the gender balance → dependence of poorer households on migrational labour revenue.

Literature References

[1] Aaberge, R., Mogstad, M. and V. Peragine (2011), “Measuring long-term inequality of opportunity”, Journal of Public Economy, 95(3-4), 193-204.

[2] Abrigo, M. R. and D. A. Ortiz (December 2018), "Devolution of Health Services, Fiscal Decentralization, and Antenatal Care in the Philippines", Philippine Institute for Development Studies, Discussion Paper Series 2018-42.

[3] Abrigo, M. R.,Tam, Z. C. and D. A. Ortiz (December 2017), "Decentralization and Health in the Philippines: A Systematic Review of Empirical Evidences", Philippine Institute for Development Studies, Discussion Paper Series 2017-58.

[4] Adler, M. D. and M. Fleurbaey (Eds.) (2016), "The Oxford Handbook of Well-Being and Public Policy", Oxford University Press.

[5] Almas, I., Cappelen, A. Lind, J. T., Sorensen, E. O. and B. Tungodden (August 2011), “Measuring unfair (in)equality”, Journal of Public Economics, 95(7-8), 488-499.

[6] Anand, S. (1983), "Inequality and Poverty in Malaysia: Measurement and Decomposition", World Bank Research Publication, Oxford University Press, 9268.

[7] Anderson, S. and D. Ray (1 December 2012), "The Age Distribution of Missing Women in India", Economic & Political Weekly, XLVII Nos 47 & 48.

[8] Anderson, S. and D. Ray (2010), “Missing Women: Age and Disease”, Review of Economic Studies, 77, 1262-1300.

[9] Angrist, J. D. and J.-S. Pischke (2009), "Mostly harmless econometrics: an empiricist's companion", Princeton University Press.

[10] Arneson, R. (1989), “Equality and Equal-Opportunity for Welfare”, Philosophical Studies, 56(1), 77-93.

[11] Asadullah, M.N. and G. Yalonetzky (2012), “Inequality of Educational Opportunity in India: Changes Over Time and Across States”, World Development, 40 (6), 151-1163.

[12] Asian Development Bank (September 2018), "Macroeconomic Update: Nepal", Asian Development Bank, 6(2).

[13] Asian Development Bank Institute (2014), "ASEAN, PRC and India: The Great Transformation", Asian Development Bank Institute .

[14] Assaad, R., Hassine, N. and D. Salehi-Isfahani (2014), “Equality of Opportunity in Educational Achievement in the Middle East and North Africa”, The Journal of Economic Inequality, 12(4), 489-515.

[15] Assaad, R., Krafft, C., Roemer, J. and D. Salehi Isfahani (2017), "Inequality of Opportunity in Wages and Consumption in Egypt", Review of Income and Wealth, 00(00).

[16] Atkinson, A. (1975), "The Economics of Inequality", Oxford University Press, Oxford.

[17] Atkinson, A. B. and A. Brandolini (2010), "On Analyzing the World Distribution of Income", The World Bank Economic Review, 1-37.

[18] Atkinson, A.B. (1970), "On the measurement of inequality", Journal of Economic Theory, 2, 244-63.

[19] Bahamonde-Birke, F. J., KUnert, U. and H. Link (2015), "About Attitudes and Perceptions - Finding the Proper Way to Consider Latent Variables in Discrete Choice Models", DIW Berlin, Discussion Papers 1474.

[20] Bahnsen, U. (28 June 2018), “Natural Born Schlaumeier: Wie intelligent sind wir, bestimmen vor allem die Gene“, Die Zeit, 27, 33.

[21] Balcazar, C. F. (July 2015), "Lower Bounds on Inequality of Opportunity and Measurement Error", The World Bank, Policy Research Working Paper 7379.

[22] Banerjee, A. V. and E. Duflo (2003), “Inequality and Growth: What Can the Data Say?”, Journal of Economic Growth, 8(3), 267–99.

[23] Banerjee, A. V. and E. Duflo (2011), "Poor Economics: A Radial Rethinking of the Way to Fight Global Poverty", PublicAffairs, Perseurs Books Group.

[24] Banerjee, A.V., Benabou, R. and D. Mookherjee (2006), “Understanding poverty”, New York: Oxford University Press.

[25] Banskota, M. (8 October 2018), Kathmandu University, personal interview.

[26] Barraclough, S and A. Domike (1966), "Agrarian Structure in Seven Latin American Countries", Land Economics, 42(4), 41-94.

[27] Barros, R., Ferreira, F., Molinas, V. and J. Chanduvi (2009), "Measuring Inequality of Opportunity in Latin America and the Caribbean", Washington, DC: Palgrave Macmillan and the World Bank.

[28] Barry, B. (2005), "Why Social Justice Matters", Cambridge, Polity Press.

[29] BBC (10 January 2019), "Nepal woman and children die in banned 'menstruation hut'", BBC, online.

[30] BBC (25 August 2015), "India's Patel Community Rally Over Caste Quotas", BBC, online publication.

[31] Behrman, J., Birdsall, N., and M. Szekely (2001), "Intergenerational Mobility in Latin America", Economia, Brookings Institution Press, 2(1), 1-44.

[32] Bennett, L. (12-15 December 2005), "Gender, Caste and Ethnic Exclusion in Nepal: Following the Policy Process From Analysis to Action", Arusha Conference "New Frontiers of Social Policy", conference paper.

[33] Bennett, L., Ram Dahal, D. and P. Govindasamy (2008), "Caste, ethnic and regional identity in Nepal: further analysis of the 2006 Nepal Demographic and Health Survey", Ministry of Health and Population, Government of Nepal.

[34] Besley, T., Burgess, R., and B. Esteve-Volart (2007), "The policy origins of poverty and growth in India", In Tim Besley, & Louise Cord (Eds.) "Delivering on the promise of pro-poor growth", London, Palgrave Macmillan and the World Bank.

[35] Beteille, T., Li, Y., Mitra, P. K., Newman, J. L. and M. Rama (2015), "Addressing Inequality in South Asia", The World Bank.

[36] Bhattachan, K.B. (2003), "Indigenous Nationalities & Minorities in Nepal", Minority Rights Group International, London.

[37] Bhattarai, D. (5 October 2018), National Natural Resource and Fiscal Commission, personal interview.

[38] Björklund, A. and M. Jäntti (2009), "Intergenerational income mobility and the role of family background" , In "Oxford Handbook of Economic Inequality", edited by Nolan, B., Smeeding, T. and W. Salverda, New York: Oxford University Press.

[39] Blackorby, C., Donaldson, D., and Auersperg, M. (1981), "A new procedure for the measurement of inequality within and between population subgroups", Canadian Journal of Economics, 14, 665-85.

[40] Bourguignon, F. (1979), "Decomposable Income Inequality Measures", Econometrica, 47(4), 901-920.

[41] Bourguignon, F. (2015), "The Globalization of Inequality", Princeton University Press.

[42] Bourguignon, F. and C. Morrisson (2002), "Inequality Among World Citizens", American Economic Review, 92, 727-744.

[43] Bourguignon, F. Ferreira, F.H.G. and M. Menendez (2007), "Inequality of Opportunity in Brazil", Review of Income and Wealth , 53 (4), 585-618.

[44] Bourguignon, F., F. Ferreira, B. Milanovic and M. Ravallion (2007), "Global Inequality and Poverty", Princeton Encyclopedia of the World Economy.

[45] Bourguignon, F., Ferreira, F. H. G. and M. Walton (2007), "Equity, Efficiency and Inequality Traps: A Research Agenda", Journal of Economic Inequality, 5(2), 235-256.

[46] Bowles, N. (18 May 2018), "Jordan Peterson, Custodian of the Patriarchy", The New York Times.

[47] Bricard, D. Jusot, F., Trannoy A. and S. Tubeuf (2013), "Inequality of opportunities in health and the principle of natural reward: evidence from European countries", Part III in "Health and Inequality, Research on Economic Inequality", 21, 335-370.

[48] Bruenig, E. (18 March 2018), "Democrats can't expand opportunity without reducing inequality", The Washington Post, Opinions.

[49] Brunori, P. (2016), "The Perception of Inequality of Opportunity in Europe" , Review of Income and Wealth, 63(3).

[50] Brunori, P., Ferreira F. Lugo, M. A. and V. Peragine (December 2013), "Opportunity-Sensitive Poverty Measurement", The World Bank, Policy Research Working Paper 6728.

[51] Brunori, P., Ferreira, F. and V. Peragine (January 2013), "Inequality of Opportunity, Income Inequality and Economic Mobility: Some International Comparison", The World Bank, Policy Research Working Paper 6304.

[52] Brzezinski (2019), "The evolution of inequality of opportunity in Europe", Applied Economics Letters, online publication.

[53] Buhmann, B., Lee, R., Guenther, S., and T. M. Smeeding (1988), "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries using the Luxembourg Income Study Database", Review of Income and Wealth , 34(1), 115-142.

[54] Buvinic, M. and G. R. Gupta (1997), "Female Headed Households and Female-Maintained Families: Are They Worth Targeting to Reduce Poverty in Developing Countries?", Economic Development and Cultural Change, 45 (2), 259 - 280.

[55] Calo-Blanco, A. and I. Garcia-Perez (2014), "On the welfare loss caused by inequality of opportunity", Journal of Economic Inequality, 12(2), 221-237.

[56] Carletto, G., Covarrubias, K. Davis, B., Krausova, M. and P. Winters (2007), "Rural Income Generating Activities Study: Methodological note on the construction of income aggregates", Food and Agricultural Organization, Guidelines.

[57] Cavalieri, M. and L. Ferrante (2016), "Does fiscal decentralization improve health outcomes? Evidence from infant mortality in Italy", Social Science & Medicine, 164, 74-88.

[58] Central Bureau of Statistics (1996), "Nepal Living Standard Survey 1995/96", Central Bureau of Statistics, National Planning Commission Secretariat , Government of Nepal.

[59] Central Bureau of Statistics (2003), "Administrative Boundaries", International Centre for Integrated Mountain Development, Government of Nepal.

[60] Central Bureau of Statistics (2004), "Nepal Living Standard Survey 2003/04", Central Bureau of Statistics, National Planning Commission Secretariat , Government of Nepal, Volumes 1 and 2.

[61] Central Bureau of Statistics (2010), "Household Questionnaire of the Nepal Living Standards Survey, 2010/11", Central Bureau of Statistics, Government of Nepal.

[62] Central Bureau of Statistics (2010), "Nepal Living Standards Survey III: Interviewer's Manual", Central Bureau of Statistics, Government of Nepal.

[63] Central Bureau of Statistics (2011), "Nepal Living Standard Survey - III 2010 : Data Dictionary" , Central Bureau of Statistics, Government of Nepal.

[64] Central Bureau of Statistics (2011), "Nepal Living Standard Survey 2010/11", Central Bureau of Statistics, National Planning Commission Secretariat , Government of Nepal, Volumes 1 and 2.

[65] Central Bureau of Statistics (2012), "National Population and Housing Census – 2011", National Planning Commission Secretariat , Government of Nepal.

[66] Central Bureau of Statistics (2012), "Nepal in Figures 2012", Central Bureau of Statistics, Nepal.

[67] Central Bureau of Statistics (2012), "Nepal Living Standards Survey – III 2010, Sampling design and implementation", Central Bureau of Statistics, Nepal.

[68] Central Bureau of Statistics (November 2012), "National Population and Housing Census 2011: National Report", Government of Nepal, National Planning Commission Secretariat, Volume 01, NPHC 2011.

[69] Central Bureau of Statistics, International Centre for Integrated Mountain Development (2003), "Districts of Nepal: Indicators of Development", Central Bureau of Statistics, Nepal.

[70] Central Bureau of Statistics, International Centre for Integrated Mountain Development (2003), "Districts of Nepal: Indicators of Development", International Centre for Integrated Mountain Development, Nepal.

[71] Chant, S. (1997), "Women-Headed Households: Diversity and Dynamics in the Developing World", London: Macmillan Press.

[72] Chavez Juarez, F. W. (2015), "Intergenerational transmission of education: the relative importance of transmission channels", Latin American Economic Review, 24(1).

[73] Chavez Juarez, F. W. and I. Soloaga (2014), "Iop- estimating ex-ante inequality of opportunity", The Stata Journal, 14(4), 830-846.

[74] Checchi, D. and V. Peragine (2010), "Inequality of Opportunity in Italy", Journal of Economic Inequality, 8(4), 429-450.

[75] Checchi, D., Peragine, V. and L. Serlenga (2010), "Fair and unfair income inequalities in Europe", IZA Discussion paper, 5025/2010.

[76] Checchi, D., Peragine, V. and L. Serlenga (2016), "Inequality of Opportunity in Europe: Is There a Role for Institutions?", Ch. 1 in Polachek, S. Tatsiramos, K. and l. Cappellari (eds.), "Inequality: Causes and Consequences".

[77] Chen, J. (2012), "An Analysis of Policy Effect on Equality of Opportunity for Health", PhD Dissertation, Yale University.

[78] Chen, J. (unpublished), "The Effect of Education policies on Equal Opportunity for Health: Evidence from U.K. and U.S.", National University of Singapore, Unpublished Manuscript.

[79] Chen, J. (unpublished), "The Equality of Opportunity for Health in the United States: An Analysis using National Longitudinal Survey of Young", National University of Singapore, Unpublished Manuscript.

[80] Chetty, R., Hendren, N., Kline, P. , Emmanuel s., and N. Turner (2014), "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility" , American Economic Review, 104(5), 141–47.

[81] Chetty, R., Hendren, N., Jones, M. R. and A. R. Porter (March 2018), "Race and Economic Opportunity in the United States: An Intergenerational Perspective", NBER, Working Paper 24441.

[82] Clark, G. (2014), "The Son Also Rises: Surnames and the History of Social Mobility", Princeton University Press.

[83] Cogneau, D. and J. Gignoux (2009), "Earnings inequalities and educational mobility in Brazil over two decades", Ch. 2 in Klasen, S. and F. Nowak-Lehman (eds.) "Poverty, Inequality and Policy in Latin America"., Cambridge, MA: MIT Press.

[84] Cogneau, D. and S. Mesple-Somps (2008), "Inequality of opportunity for income in five countries of Africa", Research on Economic Inequality, 16, 99-128.

[85] Cohen, G. A. (1989), "On the currency of egalitarian justice", Ethics, 906-944.

[86] Coller, P. (2007), "The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It", Oxford University Press.

[87] Collins, C. (2 April 2018), "Is Inequality in America Irreversible?", inequality.org.

[88] Cooper, R. J. and F. Rehman (2017), "Consumer Expectations: A Residual Based Approach", Review of Income and Wealth, 63, 841- 866.

[89] Corak, M. (2013), "Income Inequality, Equality of Opportunity, and Intergenerational Mobility", Journal of Economic Perspectives, 27(3), 79-102.

[90] Coulombe, H. and A. McKay (2008), "The Estimation of Components of Household Incomes and Expenditures: A Methodological Guide based on the last three rounds of the Ghana Living Standards Survey, 1991/1992, 1998/1999 and 2005/2006", Ghana Statistical Service, Government of Ghana.

[91] Coulter, G., Cowell, F. and S. Jenkins (1992), "Equivalence Scale Relativities and the Extent of Inequality and Poverty", Economic Journal, 102(414), 1067-1082.

[92] Cowell F. A. (2000), "Measurement of Inequality", in "Handbook of Income Distribution" edt. by Atkinson, A. B. and F. Bourguignon, 1, Ch. 2.

[93] Cowell, F. A. (1995), "Measuring Inequality", Prentice Hall / Harvester Wheatsheaf, Hemel Hempstead, 2nd edition.

[94] Cowell, F. and E. Flachaire (2014), "Statistical Methods for Distributional Analysis", Halshs, -01115996.

[95] Cowell, F.A. (2000), "Measurement of inequality", in: Atkinson, A.B., Bourguignon, F. (eds., 2000) "Handbook of Income Distribution", North Holland, Amsterdam.

[96] Cowell, F.A. (1980), "On the structure of additive inequality measures", Review of Economic Study,, 47, 521-531.

[97] Cowell, F.A. and K. Kuga (1981), "Additivity and the entropy concept: An axiomatic approach to inequality measurement", Journal of Economic Theory, 25, 131-143.

[98] Cowell, F.A., and S. P. Jenkins (1995), "How much inequality can we explain? A methodology and an application to the United States", Economic Journal, 105(429), 421-430.

[99] Cuesta, J. (2014), "Social Spending, Distribution, and Equality of Opportunities: The Opportunity Indices Analysis", World Development , 62, 106-124.

[100] Dabalen, A., Narayan, A., Saavedra-Chanduvi, J. and A. H. Suarez (2015), "Do African children have an equal chance? A human opportunity report for Sub-Saharan Africa", The World Bank, Directions in Development.

[101] Dahan, M., and A. Gaviria (2001), "Sibling correlations and intergenerational mobility in Latin America", Economic Development and Cultural Change, 49(3), 537–554.

[102] Datt, G., Jolliffe, D., Redaelli, S. and M. Shrestha (2011), "Poverty in Nepal 2010/11: estimates based on NLSS III", The World Bank, unpublished manuscript.

[103] David, P. (2007), "Path Dependence: A Foundational Concept for Historical Social Science", Cliometrica - The Journal of Historical Economics and Econometric History, 1(2), pp. 91 - 114.

[104] De Novi, C, Piacenza, M, Robone, S. and G. Turati (2019), "Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy", Regional Science and Urban Economics, 78, 103465.

[105] De Vos, G. and H. Wagamatsu (1966), "Japan's invisible race: caste in culture and personality", University of California Press.

[106] Dearden, L., Machin, S., and H. Reed (1997), "Intergenerational Mobility in Britain", The Economic Journal, 107(1), 47-66.

[107] Deaton, A. (1997), "The Analysis of Household Surveys: A Microeconometric Approach to Development Policy", John Hopkins University Press, Baltimore.

[108] Deaton, A. (2013), "The Great Escape: health, wealth and the origins of inequality", Princeton University Press.

[109] Deaton, A. and M. Grosh (1999), "Consumption", Chapter 17 in M. Grosh and P. Glewwe, (eds., 1999), "Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience", The World Bank (forthcoming).

[110] Deaton, A. and S. Zaidi (2002), "Guidelines for Constructing Consumption Aggregates for Welfare Analysis", The World Bank.

[111] Deininger, K. and L. Squire (1988), "New ways of looking at old issues: inequality and growth", Journal of Development Economics, 57, 259-287.

[112] Dhungel, N. (May 2018), "Nepal Budget for Fiscal year 2018/19", Nepal Economic Forum, online publication.

[113] Di Ciccio T. J. and B. Efron (1996), "Bootstrap confidence intervals (with Discussion)", Statistical Science, 11, 189-228.

[114] Dowding, K. (2009), "What is welfare and how can we measure it?", Oxford University Press, Ch. 19 in "The Oxford Handbook of Philosophy of Economics", edt. by Kincaid H. and D. Ross.

[115] EBRD (2017), "Transition Report 2016-17. Transition For All: Equal Opportunities in an Unequal World", European Bank for Reconstruction and Development, online.

[116] EducateNepal (01.12.2010), "Education System in Nepal", EducateNepal.

[117] Elbers, C., Lanjouw, P., Mistiaen, J.A., and B. Özler (2008), "Reinterpreting between-group inequality", Journal of Ecomomic Inequality, 6, 231-245.

[118] Elbers, C., Lanjouw, P., Mistiaen, J.A., Özler, B. and K. Simler (2004), "On the unequal inequality of poor communities", World Bank Economic Review, 18(3), 401-421.

[119] Election Commission (2013), "Gender and Inclusion Policy 2013. Gender and Inclusion Strategy 2015 - 2020", Election Commission , Government of Nepal.

[120] Emran, M. S. and F. Shilpi (2015), "Gender, Geography, and Generations: Intergenerational Educational Mobility in Post-Reform India", World Development, 72, 362–380.

[121] Emran, M. S. and F. Shilpi (Spring 2011), "Intergenerational Occupational Mobility in Rural Economy: Evidence from Nepal and Vietnam", The Journal of Human Resources, 46(2), 427-458.

[122] Emran, M. S., Greene, W. and F. Shilpi (30 July 2015), "When measure matters: coresident sample selection bias in estimating intergenerational mobility in developing countries", University Library of Munich, MPRA Paper.

[123] Esping-Andersen,G. and J. Myles (2009), "Economic Inequality and the Welfare State", Oxford University Press, Ch. 25 in "The Oxford Handbook of Economic Inequality" edited by Salverda, W., Nolan, B. and T. M. Smeeding.

[124] Ezemenari, K. M. and N. K. Joshi (June 2019), "Nepal Development Update : Investing in People to Close the Human Capital Gap (English)" , World Bank Group, Washington, D.C.

[125] Fafchamps, M. and F. Shilpi (2003), "The Spatial Division of Labour in Nepal", The Journal of Development Studies, 39(6), 23-66.

[126] Fafchamps, M. and F. Shilpi (August 2003), "The Spatial Division of Labour in Nepal", Journal of Development Studies, 39(6), 23-66.

[127] Fairchild, F.R. (1940), "Economics", Rev. ed. edn. New York, The Macmillan Company.

[128] Federal Bureau of Statistics (1991), "Pakistan – Integrated Household Survey 1991", FBS.

[129] Ferreira, F. and J. Gignoux (2014), "The Measurement of Educational Inequality: Achievement and Opportunity", The World Bank Economic Review, 28(2), 210–246.

[130] Ferreira, F. and V. Peragine (March 2015), "Equality of Opportunity: Theory and Evidence", The World Bank Group, Policy Research Working Paper 7217.

[131] Ferreira, F. H. G. and J. Gignoux (2011), "The Measurement of Inequality of Opportunity: Theory and an Application to Latin America",, Review of Income and Wealth, 57(4), 622-657.

[132] Ferreira, F. H. G. and J. Gignoux (May 2014), "The Measurement of Educational Inequality- Achievement and Opportunity", World Bank Economic Review, 28(2), 210-246.

[133] Ferreira, F. H. G. and M. Menendez (December 2007), "Inequality of Opportunity in Brazil", Review of Income and Wealth, 53(4).

[134] Ferreira, F., Gignoux, J. and M. Aran (February 2011), "Measuring Inequality of Opportunity with Imperfect Data: The Case of Turkey", The Journal of Economic Inequality, 9(4), 651-680.

[135] Ferrira, F. and V. Peragine (2016), "Individual responsibility and equality of opportunity" , Part iV in Adles, M. and M. Fleurbaey (eds) "Handbook of Well-Being and Public Policy", Oxford University Press.

[136] Fessler, P., Kasy, M. and P. Lindner (2018), "Survey mode effects on income inequality measurement", The Journal of Economic Inequality, 4, 487-505.

[137] Fields, G. (1980), "Poverty, Inequality and Development", Cambridge University Press.

[138] Fields, G. (2001), "Distribution and Development, A New Look at the Developing World", Russel Sage Foundation and MIT Press, Cambridge.

[139] Filzmoser, P., Gussenbauer J. and M. Templ (April 2016), "Deterring Outliers in Household Consumption Survey Data", Vienna University of Technology and The World Bank, Deliverable 4. Final Report.

[140] Fleurbaey, M. and V. Peragine (2013), “Ex-Ante versus Ex-Post Equality of Opportunity”, Economica, Working Paper No. 14180, 118-130.

[141] Fleurbaey, M. Peragine, V. and X. Ramos (2017), "Ex-post inequality of opportunity comparisons", Social Choice and Welfare, 49, 577-603.

[142] Food and Agriculture Organization (2008a), "Components of the Income Aggregate: 'Living Standards Survey, Nepal 2003' ", Food and Agriculture Organization.

[143] Food and Agriculture Organization (2008b), "Components of the Income Aggregate: 'Living Standards Survey, Nepal 1996' ", Food and Agriculture Organization.

[144] Food and Agriculture Organization (2013), "Components of Income Aggregate: National Survey on Household Living Conditions and Agriculture, Niger 2011", Food and Agriculture Organization.

[145] Foster J., and A.K. Sen (1997), "On economic inequality", Expanded edition. Clarendon, Oxford.

[146] Foster, J. and A. Shneyerov (2000), "Path independent inequality measures", Journal of Economic Theory, 91, 199-222.

[147] Foxley, A. (1976), "Income Distribution in Latin America", Cambridge University Press, Cambridge.

[148] Friedman, M. (1953), "Choice, Chance and the Personal Distribution of Income", Journal of Political Economy, 61(4), 277-290.

[149] Frost, M. D., Puri, M. and P. R. A. Hinde (2013), “Falling sex ratios and emerging evidence of sex-selective abortion in Nepal: evidence from nationally representative survey data”, Global Health.

[150] Furuta, H. (2016), “Users‘ Manual for Handling Resampled Micro Data of Nepal Living Standard Survey, NLSS II, 2003/04“, The Institute of Statistical Mathematics and the Statistical Information Institute for Consulting and Analysis, Version 2.0.

[151] Fuwa, N. (2000), “A Note on the Analysis of Female Headed Households in Developing Countries“, published in: Technical Bulletin of Faculty of Horticulture Chiba University, 54, 125-138.

[152] Genetti, C. (2007), “A Grammar of Dolakha Newar”, Walter de Gruyter.

[153] Georgescu, P. (11 July 2018), "Business Can Help Solve Opportunity Inequality", Forbes Magazine.

[154] Ghana Statistical Service (1995), "Ghana Living Standards Survey Report on the Third Round (GLSS3)", Ghana Statistical Service, Government of Ghana.

[155] Ghimire, B. (28 September 2018), "Privatization in Education: Private Schools Dominate National Education", The Kathmandu Post, online.

[156] GIZ (2013), "GESI Guidelines: Gender Equality and Social Inclusion Mainstreaming in INCLUDE Programme", Deutsch Gesellschaft für Internationale Zusammenarbeit GmbH, (GIZ).

[157] Glyn, A. (2011), "Functional Distribution and Inequality", Oxford University Press, in "The Oxford Handbook of Economic Inequality" edt by Nolan, B., Salverda, W., and T. M. Smeeding.

[158] Golley, J. (8 March 2018), "Do Chinese women get half the credit for holding up half the sky?", East Asia Forum, online source.

[159] Golley, J. (15 July 2015), “Unequal Opportunities, Unequal Outcomes”, China Policy Institute: Analysis, CPIAnalysis.

[160] GoNepal: True Blend of Culture & Nature (), “Caste System of Nepal”, GoNepal.eu Online.

[161] Government of the Federal Republic of Germany (2013), “Deutschalnds Zukunft gestalten, Koalitionsvertrag zwischen CDU, CSU und SPD, 18. Legislaturperiode”, CDU, CSU and SPD, government parties.

[162] Green, F., A. Dickerson, and J. S. Arbache (2001), "A Picture of Wage Inequality and the Allocation of Labor through a Period of Trade Liberalization: The Case of Brazil", World Development, 29(11), 1923-1939.

[163] Groot, L., Van der Linde, D. and C. Vindent (2018), "Inequality of opportunity in the United Kingdom, 1991 - 2008", Journal of Policy Modeling, 40, 1255 - 1271.

[164] Grosh, M. and P. Glewwe (1998), “The World Bank’s Living Standards Measurement Study Household Surveys”, Journal of Economic Perspectives, 12(1), 187-196.

[165] Grosh, M. and P. Glewwe (May 2000a), “Designing Household Surveys: Questionnaires for Developing Countries. Lessons from 15 years of the Living Standards Measurement Study”, The World Bank, Volume 1.

[166] Grosh, M. and P. Glewwe (May 2000b), “Designing Household Surveys: Questionnaires for Developing Countries. Lessons from 15 years of the Living Standards Measurement Study”, The World Bank, Volume 2.

[167] Grosh, M. and P. Glewwe (May 2000c), “Designing Household Surveys- Questionnaires for Developing Countries. Lessons from 15 years of the Living Standards Measurement Study”, The World Bank, Volume 3.

[168] Gyatso, T. (15 October 2018), Tibetan Encounters, personal interview.

[169] Handa, S. (1994), "Gender, Headship and Intrahousehold Resources", World Development, 22(10), 1535 - 1547.

[170] Hank, R. (3 September 2017), Utopien der Machtlosigkeit lassen die Menscheit seit den Anfängen nicht mehr los, F.A.S., Wirtschaft.

[171] Hassine, N. B. (2011), “Inequality of Opportunity in Egypt”, The World Bank Economic Review, 26(2), 265-295.

[172] Hassine, N. B. and A. Zeufack (May 2015), "Inequality of Outcomes and Inequality of Opportunity in Tanzania", The World Bank, Policy Research Working Paper 7260.

[173] Hastings, O. P., LaBriola, J. and D. Schneider (2018), "Income Inequality and Class Divides in Parental Investments", American Sociological Review, 83(3), 475-507.

[174] Haughton, J. H. (2009), “Handbook on Poverty and Inequality”, edited by R. S. Khandker. Washington, DC, The World Bank.

[175] Heckman, J. (1976), “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models”, Annuals of Economic and Social Measurement, 5, 475-92.

[176] Hederos, K., Jäntti, M. and L. Lindahl (2017), "Gender and inequality of opportunity in Sweden", Social Choice and Welfare, 49(3-4), 605-635.

[177] Hild, M. and A. Voorhoeve (December 2001), "Roemer on Equality of Opportunity", California Institute of Technology, Social Science Working Paper 1128.

[178] Hirsch, W. Z. (September 1961), “Income Elasticity of Public Education”, International Economic Review, 2(3), XX.

[179] Höfer, A. (1979), "The Caste Hierarchy and The State in Nepal", Universitätsverlag Wagner, University Press.

[180] Howell, D. L. (2005), “Geographies of identity in nineteenth-century Japan“, University of California Press.

[181] Hufe, P., Kanbur, R. and A. Peichl (June 2018), "Measuring Unfair Inequality- Reconciling Equality of Opportunity and Freedom from Poverty", IZA, Discussion Paper No. 11601.

[182] Ibarra, G. L. and A. L. Matines Cruz (October 2015), "Exploring the Sources of Downward Bias in Measuring Inequality of Opportunity", The World Bank, Policy Research Working Paper 7458.

[183] Insitute of Human Rights Communication Nepal (January 2008), "Successful Experiences & Lessons Learnt", Institute of Human Rights Communication Nepal.

[184] International Labor Organization (October), "Resolution concerning statistics of the economically active population, employment, unemployment, and underemployment, adopted by the Thirteenth International Conference of Labor Statisticians", ILO.

[185] International Labour Organization (2003), "Resolution concerning household income and expenditure statistics", International Labour Organization, ILO Resolution.

[186] Jantti, M. and S. Jenkins, (), , , (2015), "Income Mobility" , in "Handbook of Income Distribution" edt. by A. Atkinson and F. Bourguignon, 2, 807-935.

[187] Jenkins, S. and P. Van Kerm (2009), "The Measurement of Economic Inequality", Oxford University Press, Ch. 3 in "The Oxford Handbook of Economic Inequality" edited by Salverda, W., Nolan, B. and T. M. Smeeding.

[188] Jenkins, s. p. (2008), "Estimation and interpretation of measures of inequality, poverty, and social welfare using Stata", North American Stata Users' Group Meetings 2006, Stata Users Group.

[189] Jikdol, B. T. (7 October 2018), Kopan Monestary, personal interview.

[190] Jimenez-Rubio, D. and P. Garcia-Gomez (2017), "Decentralization of health care systems and health outcomes: Evidence from a natural experiment", Social Science & Medicine, 188, 69-81.

[191] Jodhka, S. and G. Shah (27 November 2010), " Comparative Contexts of Discrimination: Caste and Untouchability in South Asia", Economic and Political Weekly, 45(48), 99-106.

[192] Jusot, F., Tubeuf, S. and A. Trannoy (2013), "Circumstances and efforts: how important is their correlation for the measurement of inequality of opportunity in health?", Health Economics, 22, 1470-1495.

[193] Kanbur, R. (2000), "Income distribution and development", in "Handbook of Income Distribution" Atkinson, A.B., Bourguignon, F. (eds., 2000), North Holland, Amsterdam.

[194] Kanbur, R. and A. Wagstaff (July 2014), How Useful Is Inequality of Opportunity as a Policy Construct?, The World Bank Group, Policy Research Working Paper 6980.

[195] Kelekar, U. and G. M. Llanto (March 2013), "Perspectives on Health Decentralization and Interjurisdictional Competition among Local Governments in the Philippines", Philippine Institute for Development Studies, Discussion Paper Series 2013-20.

[196] Kenworthy, L. (25 October 2017), “How countries achieve economic security and equality”, Policy Options Politiques.

[197] Kenworthy, L. (May 2015), “Equality of Opportunity”, The Good Society, WordPress.org blog.

[198] Kim, K. and D. Salazar (December 2018), "Toward Gender-Focused Governance Reform in Asia", ADBI, No. 2018-4.

[199] Kincaid H. and D. Ross (2009), "The Oxford Handbook of Philosophy of Economics", Oxford University Press, edt. by Kincaid H. and D. Ross.

[200] Kis-Katos, K and B. S. Sjahrir (2017), "The impact of fiscal and political decentralization on local public investment in Indonesia", Journal of Comparative Economics, 45, 344-365.

[201] Kondor, Y. (1975), "Value Judgments Implied by the Use of Various Measures of Income Inequality", Review of Income and Wealth, 309-321.

[202] KOPILA Nepal (2018), "KOPILA Nepal Annual Report 2017/18", Kopila Nepal.

[203] Krueger, D. and F. Perri (2006), "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory", The Review of Economic Studies, 73(1), 163-193.

[204] Kruse, A., Bajracharya, R. D., and D. N. Mahato (October 2013), "Nepal Development Update 2013", World Bank Group, Washington, D.C.

[205] Kuznets, S. (1955), "Economic Growth and Income Inequality", American Economic Review, 45, 1-28.

[206] Lagarde, C. (11 April 2018), "Fix the Roof While the Window of Opportunity is Open: Three Priorities for the Global Economy", University of Hong Kong, Public Speech.

[207] Lama, N. (10 October 2018), Tribhuvan University Nepal, personal interview.

[208] Lee, E., S. (1966), "A Theory of Migration", Demography, 3(1), 47-57.

[209] Lee, L. and R.P. Trost (1978), “Estimation of Some Limited Dependent Variable Models with Application to Housing Demand”, Journal of Econometrics, 8, 357-382.

[210] Lee, T. and A. Seshadri (22 August 2018), "The moral dilemma around equality of opportunity", LSE Business Review.

[211] Lee, T. and A. Seshadri (July 2018), "Economic Policy and Equality of Opportunity", The Economic Journal, 128(612), F114-F151.

[212] Lefranc, A., Pistolesi, N. and A. Trannoy (2008), “Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike?”, Review of Income and Wealth,, 54(4), 513–46.

[213] Lefranc, A., Pistolesi, N. and A. Trannoy (2009), "Equality of opportunity and luck: definitions and testable conditions, with and application to income in France", Journal of Public Economics, 93, 1189-1207.

[214] Levine, N. (1987), "Caste, State, and Ethnic Boundaries in Nepal", Journal of Asian Studies, 46(1).

[215] Madow, W. G., Olkin, I. and D. B. Rubin (1983), "Incomplete data in sample surveys", New York, Academic Press, Vol. 1-3.

[216] Marrero, G. A. and J. G. Rodriguez (2012), "Inequality of opportunity in Europe", Review of Income and Wealth, 58(4), 597-621.

[217] Marrero, G. A. and J. G. Rodriguez (2013), "Inequality of Opportunity and Growth", Journal of Development Economics, 104, 107-122.

[218] Mazumder, B. (2005), “The Apple Falls Even Closer to the Tree than We Thought: New and Revised Estimates of the Intergenerational Inheritance of Earnings”, in Bowles,s., Gintis, H., and M. O. Groves (eds, 2005), "Unequal Chances: Family Background and Economic Success", Princeton University Press, Princeton.

[219] Milazzo, A. (2018), "Why are adult women missing? Son preference and maternal survival in India", Journal of Development Eocnomics, 134, 467-484.

[220] Millsap, A. (12 December 2018), "City Governments Should Focus on Opportunity, Not Income Inequality", Forbes Magazine, Online Version.

[221] Mincer, J. (1958), "Investment in Human Capital and Personal Income Distribution", Journal of Political Economy, 66(4), 281–302.

[222] Ministry of Finance (2017), "Development Cooperation Report: Fiscal Year 2016 / 2017", Government of Nepal, Ministry of Finance.

[223] Ministry of Finance (2018), "Economic Survey 2017/18", Government of Nepal, Ministry of Finance.

[224] Ministry of Labour and Employment (September 2014), “Labour Migration for Employment: A Status Report for Nepal: 2013/2014”, Ministry of Labour and Employment, Department of Foreign Employment, Government of Nepal.

[225] Ministry of Labour and Employment (2018), "Labour Migration for Employment- A Status Report for Nepal 2015/2016 - 2016/2017", Ministry of Labour and Employment, Government of Nepal.

[226] Molina, E., Narayan, A. and J. Saavedra-Chanduvi (December 2013), "Outcomes, Opportunity and Development: Why Unequal Opportunities and Not Outcomes Hinder Economic Development", The World Bank, Policy Research Working Paper 6735.

[227] Molinas, J. R., De Barros, R. P., Saavedra, J. and M. Giugale, with Carola, L. J., Pessino, C. and A. Hasan, (2010), "Do Our Children Have A Chance? The 2010 Human Opportunity Report for Latin American and the Caribbean", The World Bank, Conference Edition.

[228] Mookherjee, D. and A. Shorrocks (1982), "A decompositions analysis of the trend in UK income inequality", Economic Journal, 92(368), 886-902.

[229] Morgan, S., Grusky, D. and G. Fields (2006), "Mobility and Inequality: Frontiers of Research in Sociology and Economics", Palo Alto, CA: Stanford University Press.

[230] Moyo, D. (2009), "Dead Aid: Why Aid Is Not Working and How There Is Another Way for Africa", Farrar, Straus and Giroux.

[231] Moyo, D. (2012a), "Der Untergang des Westens: Haben wie eine CHance in der neuen Wirtschaftsordnung?", Piper Berlag GmbH, München.

[232] Moyo, D. (2012b), "Winner Take All: China's Race For Resources and What It Means For Us", The Penguin Group.

[233] Narayan, A. N. and R. Van der Weide (2018), "Fair Progress- Economic Mobility across Generations around the World", The World Bank.

[234] National Human Rights Commission (2017), "Annual Report Synopsis 2073-2074", National Human Rights Commission.

[235] Neidhöfer, G. Serrano, J. and L. Gasparini (2018), "Educational inequality and intergenerational mobility in Latin America: A new database", Journal of Economic Development, 134, 329-349.

[236] Nepal Law Commission (2015), "Constitution of Nepal", Nepal Law Commission, Government of Nepal.

[237] O'Donnell, O., Van Doorslaer, E., Wagstaff, A. and M. Lindelow (2008), "Measurement of Living Standards", The World Bank , Ch. 6 in "Analyzing Health Equity Using Household Survey Data" by the same authors.

[238] Ockenfels, A. (4 September 2016), "Ein strenger menschenversteher: Reinhard Selten hat as Konzept des Homo oeconomicus ins Wanken gebracht", FAZ, 35.

[239] Özler, B. (2007), "Not separate, not equal: Poverty and inequality in post-apartheid South Africa", Economic Development Culture and Change, 55(3), 487-529.

[240] Ostry, J. D. (27 October 2017), "A new narrative on income inequality and growth", Policy Options Politiques.

[241] Papanek, G. (1975), "Growth, Income Distribution and Politics in Less Developed Countries", New York Praeger, in "Economic Growth in Developing Countries" edited by Y. Ramiti, 75-96. .

[242] Pearlstein, S. (1 July 2017), "Is it inequality of income we care about - or inequality of opportunity?", The Washington Post.

[243] Peragine, V. (2004), "Measuring and Implementing Equality of Opportunity for Income", Journal of Social Choice and Welfare, 22(1), 187-210.

[244] Peragine, V., Palmisano F., and P. Brunori (2011), "Economic Growth and Equality of Opportunity", ECINEQ, Society for the Study of Economic Inequality, Working Paper, 232 (1).

[245] Peragine, V., Palmisano F., and P. Brunori (May 2014), "Economic Growth and Equality of Opportunity", World Bank Economic Review, 28(2), 247-281.

[246] Pickens, J. (1 January 2005), "Attitudes and Perceptions", Ch. 3 in "Organizational Behavior in Health Care", NY: Jones and Barlett Publishing.

[247] Pignataro, G. (2012), "Equality of opportunity: policy and measurement paradigms", Journal of Economic Surveys, 24(5), 800-834.

[248] Piketty, T. (2014), "Capital in the Twenty-First Century", Belknap, Harvard.

[249] Piraino, P. (2015), "Intergenerational Earnings Mobility and Equality of Opportunity in South Africa", World Development, 67, 396-405.

[250] Pistolesi N. (2009), "Inequality of opportunity in the land of opportunities, 1968-2001", Journal of Economic Inequality, 7, 411-433.

[251] Polumbo, B. (19 July 2018), "Millennials Shouldn't Care So Much about Inequality", National Review.

[252] Quinones, E. J., O-Campos, A. P. , Rodriguez-Alas, C., Hertz, t. and P. Winters (2009), "Methodology for Creating the RIGA-L Database", Food and Agriculture Organization.

[253] Quy-Toan, D. and L. Iyer (May 2007), “Poverty, Social Division, And Conflict in Nepal”, World Bank, Policy Research Working Papers 4228.

[254] Rama, P. (July 2016), “Decomoposing Inequality of Opportunity in Immunization by Circumstances: Evidence from India”, The European Journal of Development Research, 28(3), 431-446.

[255] Ramos, X. and D. Van De Gaer (2012), "Empirical approaches to inequality of opportunity: principles, measures and evidence", ECINEQ, Working Paper 2012-259.

[256] Ramos, X. and D. Van De Gaer (2016), “Approaches to Inequality of Opportunity: Principles, Measures, and Evidence”, Journal of Economic Surveys, 30(50), 855-883.

[257] Rawls, J. (1971), "A Theory of Justice", Cambridge, MA: Harvard University Press.

[258] Ray, D. (1998), "Development Economics", Princeton University Press.

[259] Roemer J. E. (1993a), “A pragmatic theory of responsibility for the egalitarian planner”, Wiley, Philosophy & Public Affairs, 22(2), 146-166.

[260] Roemer, J. E. (1993b), “Theories of Distributive Justice”, Harvard University Press, Cambridge.

[261] Roemer, J. and A. Trannoy (2016), “Equality of Opportunity: Theory and Measurement”, Journal of Economic Literature, 54(4), 1288-1332.

[262] Roemer, J. and B. Ünveren (2017), “Dynamic Equality of Opportunity“, Economica, 84, 322-343.

[263] Roemer, J. E. (2009), "Equality, It's Justification, Nature, and Domain", Oxford University Press, Ch. 2 in "The Oxford Handbook of Economic Inequality" edited by Salverda, W., Nolan, B. and T. M. Smeeding.

[264] Roemer, J. E. (2017), On the importance of circumstances in explaining income inequality, Presses de Sciences Po, Revue Economique, 68(1), 35-56.

[265] Roemer, J. E. (1998), “Equality of Opportunity”, Harvard University Press, Cambridge, Massachusetts.

[266] Roemer, J. E. (May 2014), "Economic Development as Opportunity Equalizer", World Bank Economic Review, 28(2), 189-209.

[267] Roemer, J. E. and A. Trannoy (2015), “Equality of Opportunity”, in “Handbook of Income Distribution”, 2A, Ch. 4, 2018-300.

[268] Roemer, J. E. and A. Trannoy (2016), "Equality of Opportunity- Theory and Measurement", Journal of Economic Literature, 54(4), 1288-1332.

[269] Rosa Dias, P. (2009), "Inequality of opportunity in health: evidence from a UK cohort study", Health Economics, 18, 1057-1074.

[270] Salverda, W., Nolan, B. and T. M. Smeeding (2009), "The Oxford Handbook of Economic Inequality", Oxford University Press, edited by Salverda, W., Nolan, B. and T. M. Smeeding.

[271] Salvi, A. (2007), “An empirical approach to the measurement of equality of opportunity“, Universita degli Studi di Milano, unpublished manuscript , (in Ramos, X. and D. Van De Gaer (2012), "Empirical approaches to inequality of opportunity: Principles, measures, and evidence", ECINEQ, Working Paper 2012 – 259).

[272] Schnitzlein, D. D. (2012), “How important is cultural background for the level of intergenerational mobility”, Economics Letters, 114(3), 335–337.

[273] Schnitzlein, D. D. (2013), “Wenig Chancengleichheit in Deutschland: Familienhintergrund prägt eigenen ökonomischen Erfolg”, Deutsches Institut für Wirtschaftsforschung e. V..

[274] Scholl, S. (September 2014), “Inequality of Opportunity in Nepal: Revelation of Minority Discrimination and the Significance of Parental Education”, Journal of the Graduate School of Asia Pacific Studies, 28, 1-28.

[275] Schuetz, G., Ursprung, H. W. and L. Woessmann (December 2005), “Educational Policy and Equality of Opportunity”, CESifo Working Paper , IZA DP No. 1906.

[276] Sen, A. (1977), "On Weights and Measures: Informational Constraints in Social Welfare Analysis", Econometrica, 45(7), pp. 1539-1572.

[277] Sen, A. (1985), “Commodities and Capabilities”, North Holland, Amsterdam.

[278] Serio, M. (11 December 2017), "Desigualdad de Oportunidades Educativas en Argentina”, Archicos Analiticos de Politicas Educativas, 25(121).

[279] Sharma, S. R. (8 October 2018), , Kathmandu University, personal interview.

[280] Shorrocks, A. F. (1980), "The class of additively decomposable inequality measures", Econometrica, 48, 613-625.

[281] Shorrocks, A.F and G. Wan (2005), "Spatial decomposition of inequality", Journal of Economic Geography, 5, 59-81.

[282] Shorrocks, A.F. (1984), "Inequality decomposition by population subgroups", Econometrica, 52, 1369-1385.

[283] Shrestha, S. (5 October 2018), Deutsche Gesellschaft für Internationale Zusammenarbeit, personal interview.

[284] Silwal, B. (14 October 2018), , KOPILA Nepal, personal interview.

[285] Singh, A. (2010), "The Effect of Family Background on Individual Wages and Examination of Inequality of Opportunity in India", Journal of Labour Research, 31, 230-246.

[286] Singh, I., Squire, L. and J. Strauss (1986), "Agricultural Household Model: Extensions, Applications, and Policy", The John Hopkins University Press.

[287] Singh, M. R. (22 May 2018), "Rural vs Urban- Measuring the Inequality of Opportunity", Down To Earth.

[288] Sir, T. (6 October 2018), Nepal in Data, personal interview.

[289] Slottje, D. J. (1990), "Using grouped data for constructing inequality indices: parametric vs. non-parametric methods", Economics Letters, 32, 193-197.

[290] Smith, N. (3 April 2018), "America, Land of Equal Opportunity? Still Not There", Bloomberg, online publication.

[291] Solon, G. (1999), "Intergenerational Mobility in the Labor Market", in "Handbook of Labor Economics", 3, Ch.29, Part A, 1761-1800.

[292] Son, H.H. (2013), Inequality of Human Opportunities in Developing Asia, Asian Development Review, 30(2), 110-130.

[293] Special Economic Zone Development Committee (2017), "Special Economic Zones (SEZs) in Nepal", Ministry of Industry, Government of Nepal.

[294] Stiglitz, J. E. (9 May 2018), "The Stockholm Statement: New Strategies for Inclusive Development", Presentation held in Nairobi, published by Columbia University (online).

[295] Stiglitz, J. E. (2015), "The Great Divide: Unequal Societies and What We Can do About Them", Norton, 1st edt..

[296] Stockholm Statement Economists (2016), "Stockholm Statement: Towards a consensus on the principles of policy-making for the contemporary world", SIDA and the World Bank, online publication.

[297] Suvedi, M. Subedi, G. and G. S. Madan. (2014), "Dimensions and Implications of Privatization of Education in Nepal "The Case of Primary and Secondary Schools"", In book "Education, Privatisation and Social Justice: case studies from Africa, South Asia and South East Asia", Editors Macpherson, I., Robertson, S and G. Walford, Symposium Books Ltd, Ch. 6.

[298] Taylor, C. and M. Hudson (1972), "World Handbook of Political and Social Indicators", New Haven, Yale University Press.

[299] The Economist (12 June 2013), "Rich Rewards", The Economist, Social Mobility, online publication.

[300] The New York Times (20 September 2017), "`Yes, we need to do better': world leaders talk democracy", The New York Times, online.

[301] Ueda, A. (2009), "Intergenerational Mobility of Earnings and Income in Japan", The B.E. Journal of Economic Analysis and Policy, 9, 1–25.

[302] Unisearch PNG & Institute of National Affairs (1996), "Papua New Guinea 1996 – Household Survey", Unisearch PNG & Institute of National Affairs, Government of Papua New Guinea.

[303] United Nations Development Program and Government of Nepal National Planning Commission (2014), "Nepal Human Development Report 2014: Beyond Geography, Unlocking Human Potential", UNDP and Government of Nepal National Planning Commission.

[304] United Nations ESCAP (2018a), "Inequality of Opportunity in Asia and the Pacific: Children's Nutrition, United Nations ESCAP, Social Development Policy Papers.

[305] United Nations ESCAP (2018b), "Inequality of Opportunity in Asia and the Pacific: Clean Energy", United Nations ESCAP, Social Development Policy Papers.

[306] United Nations ESCAP (2018c), "Inequality of Opportunity in Asia and the Pacific: Decent Work", United Nations ESCAP, Social Development Policy Papers.

[307] United Nations ESCAP (2018d), "Inequality of Opportunity in Asia and the Pacific: Education", United Nations ESCAP, Social Development Policy Papers.

[308] Valente, C. (May 2014), "Education and Civil Conflict in Nepal", World Bank Economic Review, 28(2), 354-383.

[309] Van de Gaer, D., Schokkaert, E. and M. Martinez (2001), "Three Meanings of Intergenerational Mobility", Economica, 68(272), 519–37.

[310] Varian, H.R. (2010), "Intermediate microeconomics: a modern approach", W.W. Norton & Co., New York, 8th ed. edn..

[311] Vogel, A. and K. Korinek (2012), "Passing by the Girls? Remittance Allocation for Educational Expenditures and Social Inequality in Nepal's Household 2003-2004", International Migration Review, 46(1), 61-100.

[312] Voitchovsky, S. (2005), "Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution", Journal of Economic Growth, 10, 273-296.

[313] Wagstaff, a. and R. Kanbur (2015), "Inequality of opportunity: the new motherhood and apple pie?", Health Economics, 24, 1243-1247.

[314] Williamson, J. (1998), "Growth, Inequality and Globalization: Theory, History and Policy", Cambridge University Press.

[315] Wooldridge, J. M. (2010), "Econometric Analysis of Cross Section and Panel Data", The MIT Press Cambridge, Second Edition.

[316] Wooldridge, J.M. (2009), "Introductory econometrics: a modern approach", OH: South-Western Cengage Learning, 4e edn. Mason.

[317] World Bank (), World Development Indicators (WDI), The World Bank Group.

[318] World Bank (2003), "Measuring Living Standards: Household Consumption and Wealth Indices", The World Bank, Quantitative Techniques for Health Equity Analysis - Technical Note #4.

[319] World Bank (2005), "Inequality Measures", The World Bank , in "Poverty Manual, AII", Ch. 6, 95-218.

[320] World Bank (2006), "World development report on Equity and development", Oxford University Press, Washington D.C.

[321] World Bank (2016), "Moving up the ladder: poverty reduction and social mobility in Nepal", The World Bank, Povery and Equity Globaly Practice.

[322] World Bank (2018), "The Human Capital Project", The World Bank, Report.

[323] World Bank (April 2018), "Nepal Development Update: The Challenging Path Ahead", The World Bank.

[324] World Bank (June 2011), "Large-scale migration and remittance in Nepal: Issues challenges and opportunities", The World Bank Group, Report No. 55390-NP.

[325] World Bank (8 April 2019), "Implementation of the Nepal Living Standards Survey 2019/20 (NLSS IV): Terms of Reference", The World Bank Group, Activity under P163078.

[326] Yalonetzky (2012), "A dissimilarity index of multidimensional inequality of opportunity", Journal of Economic Inequality, 10, 343-373.

[327] Yamamoto, A. (12 October 2018), Cheer Japan Overseas Cooperation Volunteers, personal interview.

[328] Yitzhaki, S. and R. I. Lerman (1991), "Income stratification and income inequality", Review of Income Wealth, 37(3), 313–329.

[329] Young, M. D. (1994), "The Rise of Meritocracy", Transaction Publishers, 2. etdn.

[330] Zoch, A. (2015), "Life chances and class: estimating inequality of opportunity for children and adolescents in South Africa", Development Southern Africa, 32(1), 57-75.

Data, Figures and Media References

[331] CIA World Factbook, https://www.cia.gov/library/publications/the-world-factbook/.

[332] Human Rights Watch, https://www.hrw.org/world-report/2018/country-chapters/nepal, Human Rights Watch.

[333] Inequality.org, http://www.inequality.org.

[334] National Planning Commission Secretariat Government of Nepal, Central Bureau of Statistics, (2012). "National Population and Housing Census – 2011", Census, http://www.census.gov.np.

[335] National Planning Commission Secretariat Government of Nepal (November 2011). "Nepal Living Standard Survey 2010/11", Central Bureau of Statistics.

[336] National Planning Commission Secretariat Government of Nepal (November 2011). "Nepal Living Standard Survey 2010/11 Report", Central Bureau of Statistics, Volumes 1 and 2.

[337] Nepal Central Bureau of Statistics (2003), "Administrative Boundaries", International Centre for Integrated Mountain Development .

[338] Nepal in Data, https://nepalindata.com.

[339] Opportunity Insight (for the US), Opportunity Project, https://opportunityinsights.org

[340] Oxford English Dictionary, online, https://www.oed.com

[341] World Development Indicators (WDI), The World Bank, http://databank.worldbank.org/Data.

APPENDICES

Appendix 1: Comprehensive Summary of Literature Review: Consumption Expenditure Aggregation

Table 5.2: Comparison Table on Guidelines for the Aggregation of Consumption Expenditure

Components	Deaton and Zaidi (2002) on NLSS I	CBS (2011, Reports 1 and 2) on NLSS III	Scholl ("optimised" methodology)	Furuta (2016) on NLSS II
Food Items	**(1) Food purchased from market**: amount spent in the typical month x 12 (or number of months typically consumed) (2) **Food that is home-produced**: quantity in typical month x farmgate price x number of months typically consumed (3) F**ood received as gift or in-kind payment**: total value for a year (4) **Meals consumed outside the home:** --> Amount spent in restaurants --> Amount spent on prepared foods --> Amount spent on meals at work [here or in work-related expenditures] --> Amount spent on meals at school [here or in education expenditures] --> Amount spent on meals on vacation [here or in vacation expenditures] **Issues** --> Missing prices or unit values, first choice is price (unit value) reported by the household; if not available, use as a proxy the median – not mean – price paid by 'similar' households in the neighbourhood, subject to checks that such prices are plausible. --> Check data for outliers; miscoding or misunderstanding of units for quantities causes errors in unit values.	Includes home production (how much would have to spend if purchased on market), food purchased for HH consumption and in-kind (value of food for past 12 months). --> both "typical" month and past 7 day consumption were reported. **Typical month criteria is used for food consumption.** --> typical months are multiplied by number of months consumed to give annual value and added to in-kind production --> tobacco products are excluded from food and placed in non-food consumption.	*tobacco taken out and placed in non-food consumption. (1) Food purchased from market: NUMBER OF MONTHS PURCHASED x AMOUNT SPENT IN SUCH A TYPICAL MONTH. 2) Food that is home-produced: quantity NUMBER OF MONTHS PURCHASED x AMOUNT MARKET PRICE IN SUCH A TYPICAL MONTH. (3) Food received as gift or in-kind payment: total value for a year (4) Meals consumed outside the home: --> meals taken outside home --> bread, biscuit, noodles --> misc. other food expenditures Both typical month and past week criterion are separately aggregated.	*tobacco taken out and placed in non-food consumption. Food items --> monthly reported values are annualised. Also for home produced foods where local market value is given. --> Reported annual value of foods that are received in kind are added.
Tobacco and selected non-food	**Daily use items:** annualise the value. **Clothing and housewares**: annualise the value	Tobacco is included from food consumption items	Tobacco is included.	Tobacco is included. ***for non-food items** --> first, monthly reported figures are anualised, because they are deemed more accurate than annual ones. When noting is reported, then the annual figures are used.
Expenditure in Utilities	Such as water, gas, electricity, telephone, etc --> can be problematic because some households are subsidised and others are not. --> Different also in quality of supplied utility (e.g. water) . --> Difficult to make accurate regional and international welfare comparison by pricing of the reported expenditures	**Excluded**: firewood: due to non-uniform and non-metric units. --> water is excluded (as in NLSS I & II). **Included**: electricity, garbage collection and telephone charges. Infrequent goods (separate classification): included in utilities --> Regular: Kerosene; matches, candles, etc --> "typical month x12. --> Irregular: Coal, charcoal; cylinder gas --> included as they are	**Not included** --> water (v02_22) (amount paid for water over the past 12 months, excluding irrigation water). Include water for HH consumption, but, difficult to distinguish form data. **Included (based on DZ, 2002)** --> wood, kerosene, call, charcoal, cylinder gas, matches, candles, lighters, lanterns, etc --> garbage, electric, phone	**Excluded** --> wood (due to different metric units) --> water is excluded due to water rationing across the country (would bias welfare ranking). **Included** --> electricity, garbage collection and telephone charges. Infrequent goods (separate classification): included in utilities --> no comment
Expenditure in health	Should only be included if they have high income elasticity in relation to their transitory variance or measurement error. --> when sick = loss of welfare. Inability to measure loss of welfare. --> irregularity in health expenditure. --> difficulty to separate "necessary" from "unnecessary" health expenditures. --> some have insurance others not. Some decide to go to public institutions (measure subsidised health value) others go private. --> the higher elasticity, the stronger the argument for inclusion.	Excluded --> because health expenditure responds to health shocks. While this expenditure should enhance welfare, the health shock previously decreased it. --> also excluded in previous rounds	**Included:** --> modern medicines and health services (doctor fees, hospital charges etc). --> traditional medicines and health services --> test for high income (consumption) elasticity (the higher, the higher the argument for inclusion)	No mention of health expenditure

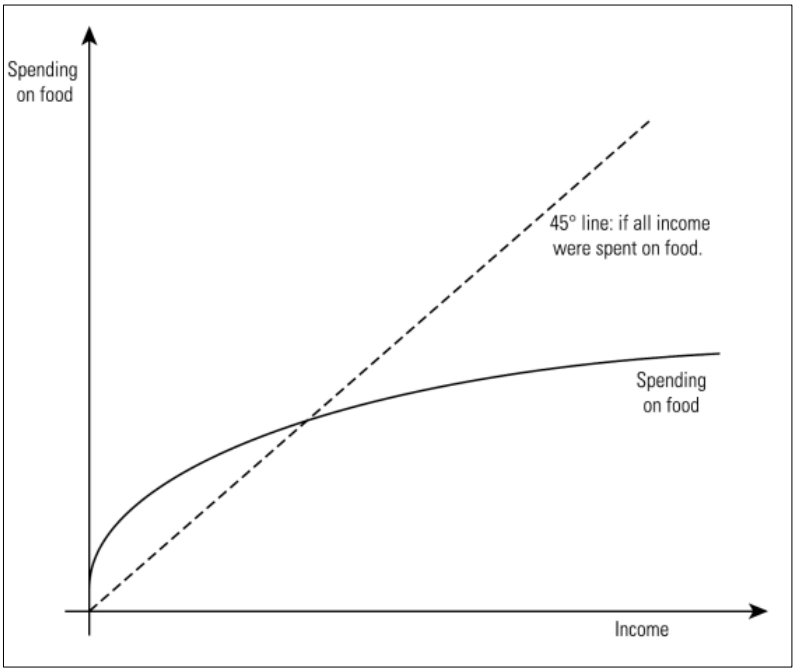
Expenditure in durable goods	**In general, for durable goods and housing:** A measure of use-value, not purchase, of durable goods is the right measure to include in the consumption aggregate from a welfare point of view. --> Exclude expenditures – instead, calculate a rental equivalent / user cost for housing & durable goods owned by the household. --> Calculate an annual rental equivalent using an appropriate real rate of interest and median depreciation values for each item calculated across all households owning that item.	**Excluded** --> some durable goods (items 412, 414-432, e.g. kitchen appliances, jewellery, electric fans, etc.): lumpy investment nature. --> But estimate of flow of service from durable goods under "durable consumption".	**Included durable goods** --> pillows, mattresses, blankets, etc; crockery, cutlery and kitchen utensils (HH use). **Own account produced** --> monetary value of good at LOCAL market over past 12 months --> rope, firewood, tailoring, shoemaking, water fencing, warehouse repairing, biogas, pickle, other, etc **Included infrequent expenses** --> legal expenses and insurances, postal expenses, fax, telephone, telegrams, excursions, holiday (including travel and lodging), car is, sports goods **Included as follows (*how many items owned*worth when acquired*depreciation rate* (depending on age, older than 0yrs).** Within past 12 months purchase: rental value without depreciation rate. --> radio, cassette, CD player, camera, bicycle, motorcycle or scooter, car, etc, a refrigerator or freezer, washing machine, fans, heaters, TV, VCR, VCD player, pressure lamps, petrol max, telephone sex, mobile phone, sewing machine, furniture, rugs, clocks, kitchen utensils, Jewellery, watches, computer, printer.	**Excluded** --> some durable goods (items 412, 414-432, e.g. kitchen appliances, jewellery, electric fans, etc.): lumpy investment nature. --> But estimate of flow of service from durable good under "durable consumption". **Included** --> estimation of durable goods of lumpy nature: purchase price (in current prices, based on the number of years ago purchased), cumulative depreciation (difference between the purchased value of the items and the current value of the items to the purchased value of the items), annual depreciation rate (median depreciation is calculated for each durable good). ***note:** average inflation rate from 2000-2003 was applied since price data was nt available at the village level.
Expenditure in education	Typically measured q**uite accurately in most surveys** -- our recommendation is to **include** them. --> Counterargument: education is an investment and not a consumption element, thus DZ argue to include in savings and not in consumption aggregate. Yet, sometimes included in consumption. --> the more children a family has, the higher educ expenditure, favouring/biasing larger HHs. Elasticity analysis. The higher the bigger the argument for inclusion.	**Excluded** --> calculated separately by combining info from non-food expenditure and parts from the education module. --> monthly fees, uniform, textbooks, etc... And scholarship. --> take maximum value of educ expenditure in educ section and the educ expenditure in "frequent non-food expenditure" section. --> shortcomings excludes individuals who have dropped out of school or are currently not going to school.	**Excluded** (Survey Section 6) --> "Education and professional services" variable in the non-food expenditure section of the survey is excluded. **Included** (Survey Section 7) --> education expenditure from the education section in the survey is included. This includes: tuition fee, other three (exams, admission, events, etc), uniform, text books and supplies, transportation, private tuition, others (snacks, tea, etc), and scholarship receive over the past 12 months.	Excluding education aggregate from section 6. --> More comprehensive education data from education section 7 is used. Values are anualised, or when only annual value is given, then that is used. --> but no information on scholarship recipients in **NLSS II.**
Consumption of housing Rental values of housing	If a household pays rent, annualise the amount of rent paid. Even if the dwelling is owned by the household or received free of charge, an estimate of the annual rental equivalent must be included in the consumption aggregate. In countries where few households pay rent, rental equivalents are potentially inaccurate, and the benefits of completeness need to be weighted against the costs of error. --> if rent has to be estimated by the dweller (e.g because is living in subsidised housing, with family, etc) then the data needs to be inspected prior use. --> missing values, for NLSS dataset, use a hedonic housing regression.	**Excluded** --> repair and maintenance and home construction improvements: investments and thus risk of double counting. **Rent or estimates** --> free housing or owned, monthly rent is estimated by dweller. --> if rented, monthly rental value. --> some unreasonable rent prices are reported (i.e. less than 100NR/month or above 30,000NR/month. --> estimation of hedonic housing regression.	Estimated rental value if own property. Estimated value if rented, squatting, provided by others (family, friends, state), other. **Unreasonable rent prices** --> Different to CBS suggestion due to exchange with CBS officials and cross-checking with other HH consumption expenditures. --> Outliers replaced by estimates of hedonic housing regression. --> New outlier definition: less than 100NR/month or above 250,000NR/month --> hedonic regression using facilities (e.g. no. of rooms, floors, roofing, etc), geographical location, access to facilities and durable assets (e.g. road, paved road). --> rent is predicted.	**Excluded** --> repair and maintenance and home construction improvements: investments and thus risk of double counting. **Rent or estimates** --> free housing or owned, monthly rent is estimated by dweller. --> if rented, monthly rental value. --> some unreasonable rent prices are reported (i.e. less than 100NR/month or above 30,000NR/month. --> estimation of hedonic housing regression.

Other	**Work-related expenses:** To the extent possible, purely work-related expenditures should be excluded. This recommendation does not include transport to work or work clothing. **Exclude:** taxes paid, purchase of assets, repayment of loans, expenditure on durable goods and housing, as well as other lumpy expenditures such as marriages and dowries. To the extent that **local property taxes** bear a relation to services rendered, we recommend their inclusion. (taxes are a deduction on income). **Leisure time:** --> Households with more leisure time have a higher level of welfare than households with no leisure. However, valuing leisure for each individual is problematic. Furthermore, it is difficult to distinguish between leisure, non-market work for the household, and involuntary unemployment. **--> Omit time and leisure in the calculation of consumption.** Gifts, charitable contributions and remittances: theory provide as much welfare as if funds used for something else. Yet, double counting problem (show up in other HHs). **Public goods:** do not include and valuation of public goods in the consumption aggregate. This is because estimating the value of e.g. access to healthcare, education, defence, police, etc is hugely problematic. HH may choose private services even if public ones are available. Some exceptions are health and housing.	**Excluded** --> taxes and fines --> marriage, funerals, dowries etc. due ot their limy nature and short time horizon.	**Excluded** --> marriage, births and other ceremonies, dowry and bride price given and received, funeral and death related expenses, expenditure on religious ceremonies, charity, cash losses, gifts and donations. --> income taxes, land taxes, housing and property taxes, the Repair and other expenses for personal vehicle, repair and maintenance of the house, repair and servicing of household effects, Home improvements and additions.	**Excluded** --> taxes and fines --> marriage, funerals, dowries etc. due to their limy nature and short time horizon
Other comments	**Data cleaning:** need to check all elements, in log and level terms for outliers. Also, graphically by using the "one-way" and "box" stata options. **Datasets:** not all datasets have the same variables → need to adjust the aggregation process. **Weights or Raising Factors:** If households interviewed in the survey had differing probability of being selected in the survey sample, household "weights" (also known as expansion a raising factors) should be included in the data. Remember to use these when deriving representative statistics for the entry under consideration. **Farm HH (with own production):** treat the farm HH as a business selling to the HH. Attempt to value produce at "farmgate" rather than "market" prices, since in developing countries, transportation costs etc can be very high. Some valuation problems include the evaluation of labour and transportation cost, quality comparability (e.g. water), etc. **Taste differences among individuals and among HHs:** Expenditure on regrettable necessities should, in theory, be excluded but in practice it is impossible reliably to distinguish between necessities and choices. Household size, however, is important and affects the household welfare associated with a given level of expenditure. --> Include expenditure on items that may or may not be regrettable necessities (e.g. work clothes). Adjust household expenditure to reflect household size. **Time reference period:** last week, typical month, over the past year. Which one to use? the one closest to a "usual consumption" that does not included outliers. DZ recommend the typical month criterion. *NLSSIII Survey (past 7 days and a "typical" month)*	**"Typical month" criterion:** --> the NLSS III (2010/11) nominal per capita consumption has been calculated on the basis of the last seven days consumption of the HH --> NLSSIII argument: inherent vagueness in the definition of "typical month" and recall bias. Advantage of past 7 day categories: less recall bias and specific time horizon, but fails to capture seasonally consumed items. --> but argued that the fieldwork across the country is divided randomly across season, this should be representative. --> NLSS I and NLSS II, past typical month or when available 12 months data was used. **Non-food expenditure** --> divided into frequent and infrequent expenditure. Regular expense items: "typical" month is multiplied by 12. Infrequent items are included as they are. **Outliers** --> CBS(2011, p. 41) 1**% of the top and 1% of the lower end of the income and consumption were excluded.**		**Major data limitations** --> Institutional HH are excluded --> Local unit of measurement was used (with conversion table) --> Data was recorded based on the memory of the respondents **Outliers** --> according to Furuta (2016, p. 165) when interviewing Nepalese delegates who conducted and analysed the survey NLSS data, **0.5% of the top and the lower end of the income and consumption were excluded.** This is because their values were deemed too far from the mean and median.

Source: CBS(2011), Furuta (2016), Deaton and Zaidi (2002)

Appendix 2: Engel's Law

Figure 5.1: Graphical Illustration of the Engel's Law



Source: Haughton (2009, p. 33)

Appendix 3: Inequality of Opportunity: Summary of Empirical Literature

Table 5.3: Main Empirical Literature Summary for IOp

Reference	Region or Country	Data source	Outcome variable	Independent variables	Method	Parametric / non-parametric	Reference Inequality of Outcome Index
Asadullah and Yalonetzky (2012)	India	NSS 1983 - 2004	Education	Parental education	GO, PC, OV	Non-parametric	Gini Index
Assaad et al. (2017)	Egypt	ELMPS 1998, 2006, 2012; LFSS 1988	Monthly individual wages, per capita household consumption	Parental education, father's occupation, birth region	Direct ex-ante	Parametric, non-parametric	MLD
Assaad et al. (2014)	Algeria, Bahrain, Dubai, Egypt, Iran, Jordan, Kuwait, Lebanon, Morocco, Oman, Palestinian National Authority, Qatar, Saudi Arabia, Syria, Tunisia, Turkey	TIMSS 1999, 2003, 2007 (students achievement in maths and sciences)	Test scores	Gender, ethnicity, parental education, number of books at home, access to computer and internet, community characteristics	Direct ex-ante	Parametric, non-parametric	Coefficient of Variation (GE(2))
Barros et al. (2009)	Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, Venezuela	Argentina: EISCA2002; Bolivia: ECH 2005; Brazil: PNAD 1995, 2005; Chile: ECSN 1996, 2006; Colombia: ECV1997, 2003; Costa Rica: EHPM 1994, 2004; Dominican Republic: ENFT 1996, 2006; Ecuador: ECV1995, 2006; El Salvador: EHPM 1998, 2005; Guatemala: ENCV 2000, 2006; Honduras: EPHPM1997, 2005; Jamaica: JSLC 1996, 2002; Mexico: ENIGH 1996, 2006; Nicaragua: ENHMNV 1998, 2005; Panama: ENV 1997, 2003; Paraguay: EPH 1999, 2005; Peru: ENH 1998, 2006; Uruguay: ENHA 2006; Venezuela: EHM 1995, 2005.	Earnings, education, housing condition (water, sanitation, electricity)	Gender, child's area of residence, HH head schooling years, per capita family income, single or two parent HH, number of siblings	HOI (D-Index)	Parametric, non-parametric, semi-parametric	-
Bourguignon et al. (2007)	Brazil	PNAD 1996	Individual earnings	Gender, race, birth region, parental education, father's occupation, (individual's: school attainment, migration, labour market status)	Indirect ex-ante	Non-parametric	Theil, Gini
Bricard et al. (2013)	Denmark, Sweden, Austria, France, Germany, Switzerland, Belgium, Netherlands, Spain, Italy, Greece, Czech Republic, Poland	Retrospective Survey of SHARELIFE 2008/09	Health	Gender, occupation, number of books at home, periods of difficulty during childhood, health status in adulthood, parent longevity, parent's health related behaviors, lifestyle/effort variables	Ex-post	non-parametric	Variance, Coefficient of Variance
Brzezinski (2019)	Austria, Belgium, Czech Republic, Germany, Estonia, Greece, Spain, Finland, France, Hungary, Ireland, Italy,	EU-SILC 2005	Disposable HH income	Gender, race, parental education, father's occupation, birth area	Direct ex-ante	Non-parametric	MLD

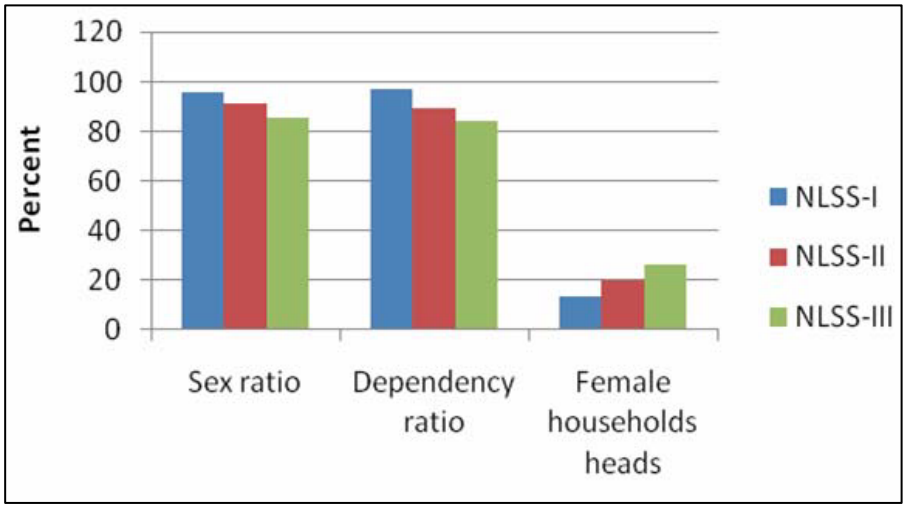
	Lithuania, Luxembourg, Latvia, Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Slovakia, UK						
Checchi and Peragine (2012)	Italy	SHIW 1993, 1995, 1998, 2000	Earnings	Parental education, gender, race	Direct ex-ante, indirect ex-ante	Parametric, non-parametric	MLD
Checchi et al. (2016)	Austria, Belgium, Bulgaria, Switzerland, Czech Republic, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Croatia, Hungary, Ireland, Italy, Lithuania, Luxembourg, Latvia, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Slovenia, UK	EU-SILC 2005	Individual disposable income	(Gender), age, country of origin, parental education	Direct ex-ante	Parametric, non-parametric	Gini Index
Cogneau and Gignoux (2009)	Brazil	PNAD 1976, 1982, 1988, 1996	Earnings (hourly), labour market, (education)	Race, birth region, father's education	Direct ex-ante, indirect ex-ante	Parametric, non-parametric, semi-parametric	Theil, Gini
Cogneau and Mesple-Somps (2008)	Ivory Coast, Ghana, Guinea, Madagascar, Uganda	Ivory Coast: EPAMCI 1985/88, Ghana: GLAA 1998, Guinea: EICVM 1994, Madagascar: EPAM 1993, Uganda: NIHS 1992	Per capita consumption	Father's education, father's occupation, birth region	Direct ex-ante	Non-parametric	Theil, Gini
Ferreira and Gignoux (2011)	Brazil, Colombia, Ecuador, Guatemala, Panama, Peru	Brazil: PNAP 1996, Colombia: ECV 2003, Ecuador: ECV 2006, Guatemala: ENCOVI 2000, Panama: ENV 2003, Peru: ENAHO 2001	Per capita consumption, per capita income	(Gender), parental education, father's occupation, ethnicity, birth region	Direct ex-ante	Parametric, non-parametric	MLD
Ferreira et al. (2011)	Turkey	TDHS 2003-2004, HBS 2003	Per capita consumption	Urban/rural, birth region, parental education, mother tongue, number of siblings	Direct ex-ante	Parametric, non-parametric	MLD, variance
Hassine (2011)	Egypt	ELMPS 2006	Monthly earning	Gender, father's occupation, parental education, birth region	Direct ex-ante, ex-post	Parametric, non-parametric	-
Hassine and Zeufack (2015)	Tanzania	NPS 2008, 2010, 2012	Income and consumption (individual and HH level)	Gender, age, parental education, parental vital status, birth place, (access to basic facilities, employment opportunities)	Indirect ex-ante, direct ex-ante	Parametric	MLD
Jusot et al. (2013)	France	ESPS since 1988	Health	Mother's health, mother's education, father's longevity	Ex-post	Non-parametric	Variance, Coefficient of Variance
Lefranc et al. (2008)	Belgium, France, West Germany, Great Britain, Italy, Netherlands, Norway, Sweden, U.S.	Belgium: PSBH 1992; France: BdF 1994; Great Britain: BHPS 1991; West Germany: GSOEP 1994; Italy: SHIW 1993; Netherlands: AVO 1995; Norway: SLL 1995; Sweden: LNU 1991; U.S.: PSID 1991.	Gross pre-fiscal annual HH income, net disposable HH income	Father's education, father's occupation	Stochastic Dominance ex-ante, GO	Non-parametric	Gini Index

Lefranc et al. (2009)	France	BdF 19790, 1984, 1989, 1994, 2000	Primary income, disposable income	HH head occupation groups, (parent's wealth)	Stochastic Dominance ex-ante, GO	Non-parametric	MLD
Marrero and Rodriguez (2012)	Austria, Belgium, Czech Republic, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Hungary, Ireland, Italy, Lithuania, Latvia, the Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Slovakia, United Kingdom	EU-SILC 2005	Equivalised personal income	Parental education, father's occupation, economic difficulties during childhood, country of birth	Direct ex-ante	Parametric	MLD
Piraino (2015)	South Africa	NIDS 2008	Individual gross income	Father's education, father's occupation, race	Direct ex-ante	Parametric, non-parametric	MLD
Pistolesi (2009)	United States	PSID 2001	Individual annual earnings	Age, parental education, father's occupation, ethnicity, birth region	Direct ex-ante	Semi-parametric	Theil, Gini, (MLD, half squared CV, std. dev. of log income)
Rosa Dias (2009)	UK	NCDS 2004	Health	Parental socio-economic background, congenital health conditions, effort (health related lifestyle e.g. smoking, fried food consumption; education attainment)	Stochastic Dominance ex-ante, GO	Non-parametric	Gini Index
Singh (2010)	India	IHDS 2004/05	Individual wage earning	Father's education, father's occupation, combination of caste and religion	Direct ex-ante	Parametric	MLD
United Nations ESCAP (2018a, b, c, d)	Afghanistan, Armenia, Bangladesh, Cambodia, India, Indonesia, Kazakhstan, Kyrgyzstan, Lao PDR, Mongolia, Pakistan, Philippines, Thailand, Turkmenistan, Viet Nam, Bhutan, Maldives, Myanmar, Tajikistan, Timor-Leste, Vanuatu	DHS, MICS (varying between 1991 to 2015)	Education, nutrition, clean water, work clean energy	Wealth, residence, education access, (gender, race)	D-Index	Parametric	Shapley Decomposition
Yalonetsky (2012)	Peru	ENAHO 2001	Level of education attainment, quality of education attained	Gender, residence, parental education, income	PC Index	Parametric, non-parametric	-
Zoch (2015)	South Africa	NIDS 2008, KIDS 2008	Education outcome	Parental education, parents presence, family income, number of children, ethnicity, residence, gender of HHH	HOI, D-Index	Parametric, non-parametric	-

Source: Author's summaries.
Note 1: PC Index = Cramer Dissimilarity Index, D-Index = Dissimilarity Index, OV = Overlap Index, GO Index = Gini Opportunity Index, HOI = Human Opportunity Index.
Note 2: Parental education indicates both mother's and father's education. Circumstantial categorical variable definitions and divisions vary between studies.

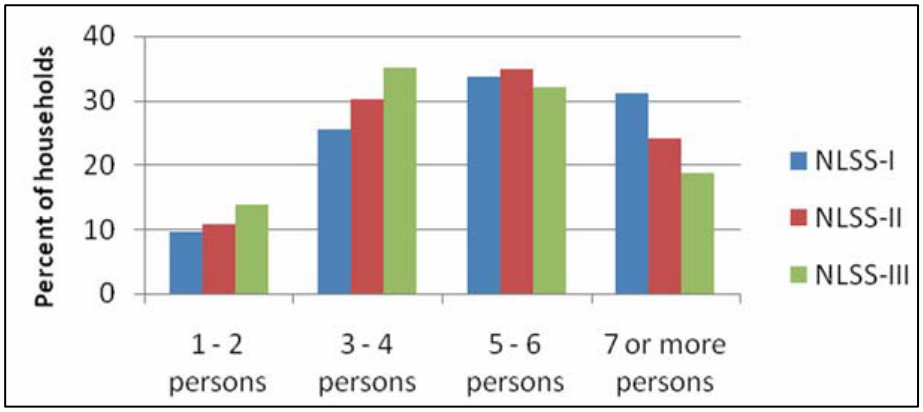
Appendix 4: Nepalese HH Characteristics and Evolution Since NLSS I

Figure 5.2: Sex Ratio, Dependency Ratio and Female HH head, 1995/6 - 2010/11 (CBS, 2011)



Source: CBS (2011, Vol. 1, p.21 – 29)

Figure 5.3: Percentage Distribution of HHs By Size, 1995/6 - 2010/11 (CBS, 2011)



Source: CBS (2011, Vol. 1, p.21 – 29)