IMPACT OF RETIREMENT ON HEALTH

A dissertation presented

by

KORYU SATO

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the subject of Economics

to

the Graduate School of Economics

at

Waseda University

December, 2023

IMPACT OF RETIREMENT ON HEALTH

BY KORYU SATO

Supervisors: Professors Haruko Noguchi and Hideo Owan

Abstract

This dissertation comprises three essays on the impacts of retirement on health. The first chapter examined the associations of retirement with cardiovascular disease and risk factors. I found a 2.2%-point decrease in the risk of heart disease and a 3.0%-point decrease in physical inactivity among retirees, compared with their working counterparts. In both sexes, retirement was associated with a decreased heart disease risk, while decreased smoking rates were observed only among women. Notably, people with high educational levels showed associations between retirement and decreased risks of stroke, obesity, and physical inactivity. Individuals who retired from non-physically demanding occupations exhibited reduced risks of heart disease, obesity, and physical inactivity, whereas those who retired from physically demanding jobs indicated an increased risk of obesity.

The second chapter investigated the impacts of retirement on cognitive function, physical independence, and self-rated health. Among men, statistically significant associations were not found, except for the realm of self-rated health, where male retirees demonstrated a 0.100 standard deviation (SD) argumentation. Conversely, female retirees showed a 0.100 SD increase in cognitive function, a 3.8%-point increase in physical independence, and a 0.193 SD increase in self-rated health concerning health outcomes. Moreover, female retirees curtailed 4.3% points in physical inactivity and 1.9% points in smoking with respect to health behaviors.

The final chapter explored the heterogeneity of retirement's impact on cognitive function using a machine-learning-based approach. The local average treatment effect indicated that retirees could recall 1.348 more words than their working counterparts. Additionally, the effects of retirement were heterogeneous, especially beneficial for women, people with higher educational attainment, elevated assets and income, those engaged in

professional clerical, or part-time occupations, those with favorable health conditions, and those frequently engaged in physical activity.

In summary, this study discerned that, on average, retirement engenders beneficial effects on health. However, these effects are heterogeneous depending on individuals' characteristics. Additionally, the findings also suggest that post-retirement health behaviors may induce the heterogeneous effects on health.

The papers on which the chapters are based are as follows; the first chapter has been published in *International Journal of Epidemiology*, titled "Retirement and Cardiovascular Disease: A Longitudinal Study in 35 Countries," coauthored by Haruko Noguchi, Kosuke Inoue, Ichiro Kawachi, and Naoki Kondo; the second chapter has been presented as a working paper in *the Social Science Research Network (SSRN)*, titled "Sex Differences in the Impact of Retirement on Health: Evidence from 35 Countries," coauthored by Haruko Noguchi; the final chapter has been presented as a working paper in *SSRN*, titled "Heterogeneous Treatment Effect of Retirement on Cognitive Function," coauthored by Haruko Noguchi and Kosuke Inoue.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2SLS	Two-stage least squares
ADL	Activities of daily living
BMI	Body mass index
CHARLS	China Health and Retirement Longitudinal Study
CI	Confidence interval
CLATO	Conditional local average treatment effect on the overlap population
Coef.	Coefficient
CRELES	Costa Rican Longevity and Healthy Aging Study
CVD	Cardiovascular disease
ELSA	English Longitudinal Study on Ageing
ERA	Early retirement age
FE	Fixed effect
FEIV	Fixed effects instrumental variable
GMM	Generalized method of moments
HRS	Health and Retirement Study
IADL	Instrumental activities of daily living
IV	Instrumental variable
JSTAR	Japanese Study of Aging and Retirement
KLoSA	Korean Longitudinal Study of Aging
LATE	Local average treatment effect
LATO	Local average treatment effect on the overlap population
Obs.	Observations
OLS	Ordinary least squares
ORA	Official retirement age
SD	Standard deviation
SHARE	Survey of Health, Ageing and Retirement in Europe
SPA	State pension age
MHAS	Mexican Health and Aging Study

Acknowledgments

Three studies constituting this dissertation were supported by the Japan Society for the Promotion of Sciences (grant number: 20K18931, 23H03164) and the Health Care Science Institute Research Grant. The funders had no role in the study design; in the collection, analysis, and interpretation of the date; in the writing of the report; or in the decision to submit the article for publication. There are no conflicts of interest to declare. This dissertation used data or information from: the Harmonized SHARE dataset and Codebook, Version F as of June 2022; the Harmonized ELSA dataset and Codebook, Version G.2 as of July 2021; the Harmonized CRELES dataset and Codebook, Version A as of August 2016; the Harmonized MHAS dataset and Codebook, Version C as of September 2022; RAND HRS Longitudinal File 2018 (V2) and the Harmonized HRS dataset and Codebook, Version C as of January 2022; the Harmonized CHARLS dataset and Codebook, Version D as of June 2021; the Harmonized JSTAR dataset and Codebook, Version B as of August 2014; and the Harmonized KLoSA dataset and Codebook, Version D.2 as of July 2021. The survey harmonization was funded by the National Institute on Aging (grant number: R01 AG030153, RC2 AG036619, R03 AG043052) and conducted by the Gateway to Global Aging Data in collaboration with the research team of the surveys. The HRS was sponsored by the National Institute on Aging (grant number: NIA U01AG009740) and was conducted by the University of Michigan. The MHAS received support from the National Institute on Aging (grant number: R01 AG018016) in the United States and the Instituto Nacional de Estadística y Geografía (INEGI) in Mexico. The harmonized datasets and more information are available through the Gateway to Global Aging Data website (https://g2aging.org/). The studies used publicly available data that obtained informed consent from all participants and ethical approval from relevant local ethics committees. Thus, the Ethics Committee of Kyoto University exempted this study from review.

I am profoundly grateful to my dissertation committee: Haruko Noguchi, Hideo Owan, and Toshiaki Iizuka for their unwavering dedication in providing invaluable feedback. Haruko Noguchi's encouraging attitude consistently left me feeling positive after our interactions. Although I had prior publications in epidemiology, venturing into economics was a new endeavor for me. I extend my heartfelt thanks to her for patiently and meticulously correcting my papers. I am also deeply appreciative of Hideo Owan, who generously shared his profound expertise. His comments consistently delved to the core of my research and significantly contributed to its refinement. Despite residing in Kyoto, he extended online facilities and warmly welcomed me to his seminars, offering a precious opportunity to stay updated on current evidence and connect with colleagues in the field of personnel economics. Choosing to pursue my doctoral degree at Waseda University, where Professors Noguchi and Owan were based, was unquestionably the correct decision. Furthermore, I extend my utmost gratitude to Toshiaki Iizuka for agreeing to serve on my committee despite his affiliation with a different university. His outstanding work in the field of health economics has always been an inspiration to me, and he remains one of my key role models. I consider it an honor to have him as a committee member.

I would like to express my sincere thanks to the coauthors of the three studies that constitute this dissertation. The study presented in the first chapter was coauthored by Haruko Noguchi, Kosuke Inoue, Ichiro Kawachi, and Naoki Kondo; the second chapter was coauthored by Haruko Noguchi; and the third chapter was coauthored by Haruko Noguchi and Kosuke Inoue. Ichiro Kawachi, my advisor at Harvard University, introduced me to the allure of research, shaping my decision to pursue this path. Naoki Kondo, my supervisor at Kyoto University, welcomed me into a wonderful research environment despite my prior role at the Ministry of Health, Labour and Welfare. Kosuke Inoue, my colleague at Kyoto University, has not only been a brilliant collaborator but also a kind and supportive friend. Our lunchtime discussions, spanning both research and personal life, were invaluable to me. Without the contributions of any one of them, my dissertation would not have reached completion. I am also grateful for the insightful comments from fellow members of Noguchi's and Owan's seminars, as well as other conference attendees.

I also extend my gratitude to my parents. Tragically, my father passed away from a heart attack due to overwork when I was just a year old. Despite the circumstances, the Ministry of Health, Labour and Welfare initially did not recognize his death as a workplace-related injury, necessitating a legal battle that lasted 18 years for compensation. Ultimately, we prevailed in court, and the compensation enabled me to pursue higher education. Motivated by this experience, I aspired to delve into health economics with the aim of contributing to the creation of a healthier society. I remain committed to publishing valuable research while treasuring the legacy my parents have left me.

Lastly, my deepest appreciation goes to my wife and two children, who have consistently stood by my side throughout this journey.

Introduction

The global population of people aged over 60 years will undergo a twofold increase between 2015 and 2050 (World Health Organization 2020). As the aging process ensues, molecular and cellular damage gradually accumulates and results in cognitive and physical impairments and morbidities (Kirkwood 2008). Accordingly, the number of years of life lost due to population aging is estimated to rise globally by the year 2040 (Foreman et al. 2018). However, the aging process is not homogeneous across populations, and some people enjoy a longer disability-free life expectancy, while others not only live shorter lives but also spend more time with impairments (World Health Organization 2015). In recognition of these divergences, the United Nations has designated the period of 2021 to 2030 as the Decade for Healthy Aging (World Health Organization 2020). The purpose of this initiative is to strategically address and mitigate health disparities within the aging demographic and enhance the overall well-being of older adults.

Various factors intricately affect older adults' health, and prevailing theories suggest that retirement may lead to health disparities in later life by exerting either a positive or negative impact. The "use it or lose it" hypothesis (Salthouse 2006; Twomey and Taylor 1984) postulates that retirement could adversely affect cognitive and physical function due to the potential loss of intellectual stimulation and physical activity inherent in the work environment. Based on the Grossman model (Grossman 1972), Mazzonna and Peracchi (2012) hypothesized that retirees might possess reduced incentives to invest in health to retain productivity within the labor market. In contrast, Atalay, Barrett, and Staneva (2019) argued that retirees could potentially allocate more time toward engaging in health-bolstering activities compared to individuals of working age. Recently, numerous developed countries have taken the initiative to elevate their state pension age (SPA) to encourage continued work participation among older adults. This strategic move serves to improve pension finances and address the challenges posed by the rapidly aging population (Organisation for Economic Co-operation and Development, 2021). However, the impacts of postponed retirement on health remain unclear, and a consensus on this matter is notably absent (Garrouste & Perdrix, 2022; Nishimura et al., 2018; Xue et al., 2020).

The discordant findings in previous studies can be attributed to potential sources of inconsistency, such as divergent statistical methodologies, varying study designs, disparate measures utilized to define retirement and health outcomes, and discrepancies in the countries under examination. Thus, comprehensive cross-country investigations utilizing consistent research methodologies are needed. This dissertation serves to fill this gap by applying a unified methodology to harmonized cross-country datasets, thus making a noteworthy contribution to the existing literature by providing a holistic view of the impact of retirement on health. Furthermore, an additional hypothesis posited in this research is that the inconsistent results in the previous literature can partly stem from the effect heterogeneity inherent in the retirement phenomenon. When a subgroup, adversely influenced by retirement, conceals its beneficial effect within other subgroups, the average treatment effect of retirement on the population becomes obscured. Therefore, classical interaction and stratification analyses were employed in conjunction with a novel machine-learning-based approach to reveal that the effects of retirement on health are heterogeneous depending on individuals' characteristics. This elucidation stands as another notable contribution of this dissertation.

The first chapter of this dissertation examined the intricate associations of retirement with cardiovascular disease (CVD) alongside its associated risk factors. Given that CVD stands as the leading cause of mortality among the older population worldwide, it warrants comprehensive investigation. The study leveraged harmonized longitudinal datasets from the Health and Retirement Study, in concert with its sister surveys spanning 35 countries. The dataset comprised 396,904 observations from 106,927 unique individuals aged 50–70 years, with a mean follow-up period of 6.7 years. To unveil insights into the causal relationship, fixed effects instrumental variable regressions were performed, using the SPA as an instrument variable for retirement. The empirical analysis yielded substantive findings, notably a 2.2%-point decrease in the risk of heart disease and a 3.0%-point decrease in physical inactivity among retirees compared with their working counterparts. This protective effect of retirement against heart disease was evident across both genders, while the reduction in smoking prevalence was observed only among women. Moreover, people with high educational attainment showed compelling associations between retirement and decreased risks of stroke, obesity, and physical inactivity. A further stratified analysis divulged that people who retired from non-physically demanding professions exhibited reduced risks of heart disease, obesity, and physical inactivity. Conversely, those who retired from physically demanding occupational domains indicated an increased risk of obesity. In summary, the deductions drawn from this chapter coalesce to affirm that retirement, on average, correlates with a reduced risk of heart disease. Additionally, some associations of retirement with CVD and risk factors appeared heterogeneous by individual characteristics.

The second chapter of this dissertation recognized that cognitive and physical impairments can occur through different pathways from CVD. Therefore, it meticulously investigated how retirement influences cognitive function, physical autonomy, and self-rated health. Using the same comprehensive datasets as the first chapter, the investigation yielded discerning results. Among men, no statistically significant association was observed, except for the domain of self-rated health, where male retirees demonstrated a 0.100 standard deviation (SD) increase. Conversely, female retirees showed substantial enhancements. Specifically, there was a 0.100 SD increase in cognitive function, a 3.8%-point increase in physical independence, and a 0.193 SD increase in self-rated health compared to pre-retirement health outcomes. Moreover, female retirees curtailed commendable reductions in physical inactivity by 4.3% points and smoking by 1.9% points with respect to health behaviors. These findings led to the conclusion that the observed disparities in post-retirement health behaviors between sexes may contribute to heterogeneous effects on health. The evidence underscored the need to consider gender-specific responses and adaptions to retirement in understanding its impacts on various health dimensions.

The final chapter of this dissertation, building upon the nuanced insights garnered from preceding chapters showcasing the diverse effects of retirement contingent upon individuals' characteristics, employed an innovative analytical approach—instrumental variable causal forests—to thoroughly explore the effect heterogeneity of retirement on cognitive function. Employing harmonized data from 19 countries, the analysis encompasses a thorough assessment of heterogeneity across a spectrum of 60 covariates. The local average treatment effect unveiled a tangible cognitive benefit for retirees, indicating their capacity to recall 1.348 more words than their working counterparts. Additionally, the effects of retirement demonstrated substantial heterogeneity, notably exhibiting pronounced benefits for specific subgroups. This included women, people with high educational attainment, substantial assets, and high income, those engaged in professional, clerical, or part-time occupations, those maintaining good health conditions, and those actively involved in frequent physical activity. Drawing on prior revelations regarding the potential acceleration of cognitive decline through delayed retirement for specific demographic segments, an estimation of monetary costs related to dementia care, induced by the increase in the SPA, was meticulously conducted. The estimation predicted that the United Kingdom would incur a more substantial financial burden than the United States, primarily attributed to the absence of early retirement options and the consequential widespread impact on the working population due to the SPA escalation. In essence, this chapter affirmed that the implications of retirement on cognitive function are multifaceted and contingent on a myriad of individual characteristics, underlining the imperative to tailor policy and healthcare initiatives accordingly.

In conclusion, this dissertation provides three crucial policy implications. First, it establishes that retirement generally contributes to better health outcomes. Policymakers should weigh the societal benefits of raising the SPA and affording older people the opportunity to remain in the workforce against the potential societal costs stemming from heightened risks of costly medical conditions such as CVD and dementia. Second, the research underscores the heterogeneous effects of retirement on health, depending on individuals' characteristics. Thus, advocating for flexibility in retirement through early retirement options within the pension systems is strongly recommended. This flexibility empowers individuals to make informed decisions regarding their retirement timing, aligning with their unique circumstances. Finally, the dissertation suggests that post-retirement health behaviors may inadvertently accentuate health disparities. Given the global trend of increasing SPA, there is an imperative to promote healthy behaviors post-retirement. This proactive approach can effectively mitigate the potential adverse health effects resulting from delayed retirement, aligning with the broader global goal of enhancing overall public health.

1. Retirement and Cardiovascular Disease: A Longitudinal Study in 35 Countries[1]

1.1 Introduction

Many countries have been increasing the state pension age (SPA) to accommodate the aging population (Organisation for Economic Co-operation and Development, 2021). For example, the United Kingdom and the United States plan to increase their SPA to age 67. The SPA influences individual workers' decisions regarding the timing of their retirement. Nonetheless, the potential impact of delayed retirement on health has not been considered in political debates. In particular, cardiovascular disease (CVD) ranks as the leading cause of mortality worldwide, killing 16.5 million people aged ≥55 in 2019 (GBD 2019 Diseases and Injuries Collaborators, 2020). A growing body of literature has explored the association between retirement and CVD; however, the findings are inconsistent. Several studies from European countries found an increased CVD risk among retirees, whereas studies from the United States seldom showed a clear association between retirement and CVD (Xue et al., 2020). No studies reported a beneficial association between retirement and CVD (Xue et al., 2020).

The observed detrimental association may be attributable to the healthy worker survivor effect ("those who remain employed tend to be healthier than those who leave employment" (Arrighi & Hertz-Picciotto, 1994)). There is conflicting evidence of an increased CVD risk in the literature. For instance, job strain is a known risk factor for CVD (Kivimäki et al., 2012). Based on the psychosocial mechanism, relief from job strain can be protective against CVD. There are other inconsistencies in several findings regarding preferable changes in CVD risk factors, such as increased physical activity, sleep quality, and smoking cessation after retirement (Barnett et al., 2012; Kämpfen & Maurer, 2016; Myllyntausta et al., 2018; Müller & Shaikh, 2018; Celidoni & Rebba, 2017; Kesavayuth et al., 2018; Eibich, 2015; Syse et al., 2017; Insler, 2014). Although evidence is mixed, some studies also suggested decreases in body weight, hypertension, diabetes, and heavy drinking among retirees (Insler, 2014; Syse et al., 2017; Xue et al., 2017). To address the potential healthy worker survivor effect, several studies have used the SPA as an instrumental variable (IV), which is strongly correlated with retirement but does not directly affect the outcomes.

[1] The results presented in this chapter have been published as Sato et al. (2023).

IV studies using data from the Health and Retirement Study (HRS) and the English Longitudinal Study on Ageing (ELSA) have reported an ambiguous association between retirement and CVD (Behncke, 2012; Coe & Lindeboom, 2008). Because IV estimates tend to have wide confidence intervals (CIs) (Glymour & Swanson, 2020), these studies may be less conclusive. Most previous studies were conducted in a single country or region and had limitations with respect to statistical power and generalisability to other countries. Moreover, researchers were unable to determine whether the inconsistent results were due to differences in the study population or other factors (e.g., study designs, measures of retirement and outcomes, and analytic methods).

Therefore, the present study aimed to investigate the association of retirement with CVD and various risk factors and provide a holistic view using data from 35 countries. The endogenous decision regarding retirement was handled using the SPA as an IV.

1.2 Methods

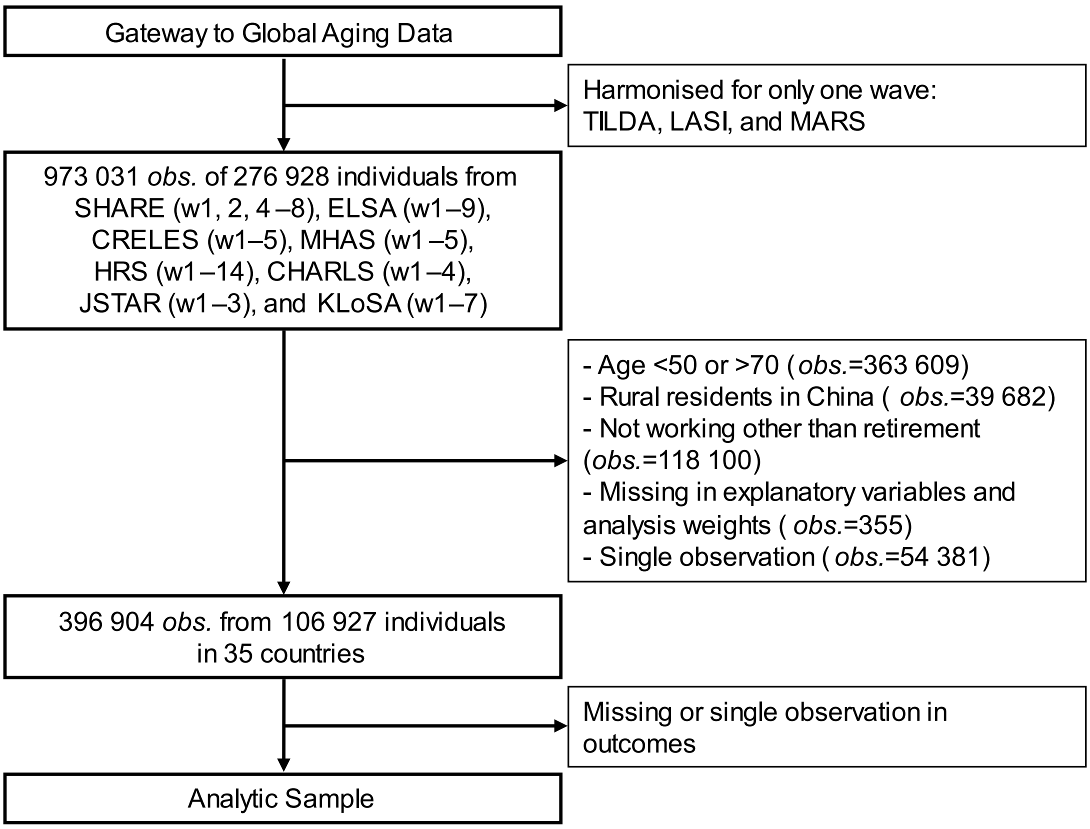
1.2.1 Database and Study Participants

This study used the harmonized datasets of the HRS and its sister surveys provided by the Gateway to Global Aging Data project (Lee et al., 2021). Our datasets comprised the following surveys with multiple observations: waves 1, 2, and 4–8 (2004–2019) of the Survey of Health, Ageing and Retirement in Europe (SHARE), waves 1–9 (2002–2018) of the ELSA, waves 1–5 (2005–2012) of the Costa Rican Longevity and Healthy Aging Study (CRELES), waves 1–5 (2001–2018) of the Mexican Health and Aging Study (MHAS), waves 1–14 (1992–2018) of the RAND HRS, waves 1–4 (2011–2018) of the China Health and Retirement Longitudinal Study (CHARLS), waves 1–3 (2007–2011) of the Japanese Study of Aging and Retirement (JSTAR), and waves 1–7 (2006–2018) of the Korean Longitudinal Study of Aging (KLoSA). All surveys were designed to represent the national older population, except for the JSTAR, which recruited participants randomly from 10 specific municipalities. The same individuals were followed up approximately biennially; however, the MHAS and CHARLS conducted the interviews triennially since 2012 and 2015, respectively. The CRELES included a cohort interviewed in waves 1–3 (2005–2009) and another cohort interviewed in waves 4–5 (2010–2012).

Figure 1.1 describes the flowchart of our analytic sample. Originally, the harmonized data involved 973,031 observations from 276,928 unique individuals. In the analysis, we included people aged 50–70 years, whose timing of retirement could be affected by the SPA. Of note, the CRELES interviewed adults aged ≥60 years in waves 1–3 and individuals aged

55–65 years in waves 4–5, whereas the HRS interviewed adults aged ≥51 years. We excluded the following observations in the analyses: rural residents in the CHARLS because China had different pension systems in rural and urban areas (Lei & Liu, 2018); those who were not working for reasons other than retirement (e.g., unemployed, disabled, homemaker); observations with missing values for explanatory variables; individuals with only one observation because keeping them in a fixed-effects model could underestimate standard errors (Correia, 2015). Thus, our study participants comprised 396,904 observations from 106,927 unique individuals in 35 countries. We additionally excluded observations with missing outcomes from the analysis, and the number of missing values varied across outcomes.

Figure 1.1 Flowchart of the Analytic Sample



TILDA stands for the Irish Longitudinal Study on Ageing; LASI stands for Longitudinal Aging Study in India; MARS stands for Malaysia Ageing and Retirement Study.

1.2.2 Outcomes

The outcomes included the occurrence of heart disease and stroke. At the first interview, participants were asked whether a doctor had ever told them that they had or currently have these conditions (see Appendix A and Hu & Lee (2012) for details). The variable indicated 1 if the participants had ever had the condition and 0 if otherwise. Their previous reports were carried forward to the subsequent waves, and they were asked about updates from the last interview. If participants later disputed reports from previous waves, they were corrected retrospectively.

We also investigated six CVD risk factors, namely hypertension, diabetes, obesity, physical inactivity, smoking, and binge drinking, although the data on health behaviors were not collected in some countries (see Appendix B for details). We hypothesized that reduced unhealthy behaviors (i.e., physical inactivity, smoking, and binge drinking) would prevent hypertension, diabetes, and obesity and result in a decreased occurrence of heart disease and stroke. Diagnosed health conditions of hypertension and diabetes were asked in the same manner as heart disease and stroke. Obesity was defined as a body mass index (BMI) of 30 kg/m^2 or higher (World Health Organization, 2021). We considered those who engaged in vigorous or moderate physical activity less than once per week as physically inactive individuals. Smoking status indicated whether the participants were currently smoking. Binge drinking was defined as consuming five or more drinks per day for men and four or more for women (Centers for Disease Control and Prevention, 2022). All variables were coded as binary.

1.2.3 Retirement Status

Retirement status was determined based on the harmonized variable of self-reported labor force status (see Appendix C and Zamarro & Lee (2012) for details). We considered individuals who mentioned retirement in an interview as retirees, irrespective of whether they were currently working (i.e., including those "partly retired"), and compared them with workers, as in the previous literature (Celidoni & Rebba, 2017; Kesavayuth et al., 2018; Müller & Shaikh, 2018). There is a narrower definition of retirement, which refers to the complete exit from the labor market (Behncke, 2012; Celidoni & Rebba, 2017; Eibich, 2015). Hence, considering this alternative definition, we determined individuals who declared being retired but were currently working by combining the variables of labor force status and engagement in paid work; subsequently, we excluded these individuals as a robustness check.

1.2.4 Instrumental Variables

We used the SPA as an IV for retirement to eliminate the potential healthy worker survivor effect. In some countries, early pension is available with reduced benefits or sufficient years of social security contributions. Thus, we used the early retirement age (ERA) and official retirement age (ORA) as joint instruments to predict retirement, as in a previous study (Eibich, 2015). A dummy variable of ERA indicated whether the participants had attained the earliest age at which individuals are entitled to reduced pensions or full pensions with some conditions. That of ORA indicated whether the participants had attained the age at which individuals are entitled to minimum guaranteed pensions or full pensions without any conditions. For countries without an early pension, the ERA variable was set to zero for all participants. For each country, we collected the data on ERA, ORA, and their changes during the study period (Appendix D). Appendix E shows a graph depicting age and the corresponding retirement rate by country; changes in the retirement rate around the SPA were observed.

1.3 Statistical Analyses

1.3.1 Empirical Model

We investigated the association of retirement with CVD and risk factors using linear probability models estimated by the fixed-effects instrumental variable (FEIV) method with the two-stage least squares procedure. In the first stage, the probability of retirement was predicted as follows:

$$R_{ijt} = \beta_1 age_{ijt} + \beta_2 age_{ijt}^2 + \beta_3 ERA_{ijt} + \beta_4 ERA * age_{ijt} + \beta_5 ERA * age_{ijt}^2 + \beta_6 ORA_{ijt} + \beta_7 ORA * age_{ijt} + \beta_8 ORA * age_{ijt}^2 + \beta_9 mstat_{ijt} + \alpha_i + u_j + \lambda_t + u_j * \lambda_t + \varepsilon_{ijt}$$

where R_{ijt} denotes whether individual *i* residing in country *j* was retired in interviewed year *t*. ERA_{ijt} and ORA_{ijt} are instruments indicating whether the participants reached the ERA and the ORA. We include age and age squared to account for a normal aging process, assuming that cognitive and physical function decline at an increasing rate with aging (Bonsang et al., 2012). Additionally, the model incorporates interactions between the age function and the IVs because the slopes of the retirement probability vary after reaching SPAs as shown in Appendix E.[2] $mstat_{ijt}$ denotes whether the participants were married. α_i, u_j, and λ_t are FEs for individuals, countries, and survey years, respectively. The model also

[2] Previous studies also included interaction terms between age and the SPA in their empirical model (Nishimura et al., 2018; Rose, 2020).

includes interaction between country and year FEs. ε_{ijt} is an error term. Subsequently, the second stage was estimated using the following equation:

$$Y_{ijt} = \gamma_1 \hat{R}_{ijt} + \gamma_2 age_{ijt} + \gamma_3 age_{ijt}^2 + \gamma_4 mstat_{ijt} + \alpha_i + u_j + \lambda_t + u_j * \lambda_t + \nu_{ijt}$$

where Y_{ijt} is health outcomes and health behaviors as risk factors, $\hat{R}_{ijt}$ is the predicted probability of retirement from the first-stage estimation, and ν_{ijt} is an error term.

The FEIV model has several advantages with respect to pooling data from different countries and estimating the potential causal retirement effect on outcomes. We included FEs of individuals and countries in the model, which controlled for both observable and unobservable time-invariant factors, such as genes and educational attainment as well as institutional and cultural differences across countries. To account for time-variant factors, we adjusted for individuals' centered age, its squared term, and marital status. Additionally, we included year FE and its interaction with country FE to capture global and country-specific time trends. To eliminate the healthy worker survivor effect, we applied the IV method using the ERA and ORA as instruments for retirement; at the first stage estimation, the retirement probability was predicted by the ERA, ORA, and their interactions with age and age squared, as in a previous study (Rose, 2020), along with other time-variant variables. We compared the FEIV estimates with those of FE models without IVs.

We also assessed heterogeneity across several subgroups. First, we checked heterogeneity across countries using I^2 statistics as an analogy to meta-analysis (Higgins & Thompson, 2002). Additionally, the test of interaction was performed across the region (i.e., Europe [including Israel], America [Costa Rica, Mexico, and the United States], and Asia [China, Japan, and South Korea]) and income level (i.e., high-income countries and low-middle income countries [Bulgaria, Romania, Costa Rica, Mexico, and China]). Second, we examined sex differences because the SPA, employment environment, and CVD risks differ by sex (Regitz-Zagrosek & Kararigas, 2017). Third, we stratified participants according to educational attainment because previous studies reported that people with higher education levels tend to present a more evident association between retirement and an increase in physical activity (Barnett et al., 2012; Celidoni & Rebba, 2017; Kämpfen & Maurer, 2016). Fourth, we stratified the participants based on whether they had experienced physical labor or not because previous studies showed that retirement from physically demanding jobs was associated with increased obesity and decreased physical activity (Chung et al., 2009; Stenholm et al., 2017).

In all analyses, individual-cluster robust standard errors were estimated. All analyses were performed using Stata version 17.0 (StataCorp, College Station, TX, USA), except for multiple imputation in robustness checks.

1.3.2 Robustness Checks

First, considering the alternative definition of retirement (i.e., fully retired), we excluded from our analysis those who reported being retired but were still working (i.e., partly retired). Second, we narrowed the age window to 52–68 years to check the robustness of the findings. Third, we excluded countries in which the IV appeared to be weakly correlated with retirement (i.e., F statistic was below the Stock-Yogo's critical value of 10% maximal relative bias (Stock & Yogo, 2002)). Fourth, given that 24.1% of the pooled sample came from the United States, we excluded data from the HRS and checked the robustness of the findings. Fifth, to reduce potential bias from missing observations, we adopted multiple imputation (Honaker & King, 2010) using R 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria). Sixth, given that outcomes were binary, we also performed FEIV Poisson regressions using the control function approach (Lin & Wooldridge, 2019). Seventh, we excluded those who retired within two years from the analysis to reduce the healthy worker survivor effect in another way. Eighth, recognizing the impact of marital status on retirement choices, we incorporated interactions between IVs and marital status into the FEIV models. Finally, as the retirement effect on CVD has been suggested by previous studies to be time-varying (Coe & Lindeboom, 2008; Moon et al., 2012), we examined the short- and long-term retirement effects. Each group of retirees who retired within five years and who retired over five years ago was compared with those who were working.

1.4 Results

1.4.1 Descriptive Statistics

A total of 106,927 individuals were followed up for a mean period of 6.7 years (Table 1.1). Some participants failed to be followed up. Nonetheless, we confirmed that there was no difference in characteristics between individuals who were followed up and those who were lost to follow-up, except for age; those who were lost to follow-up were older by 0.84 years than those who were followed up (Appendix Table F.1).

Table 1.1 Cohort Characteristics of the Surveys

Survey	Country	Interview years	No. of unique individuals	Mean follow-up period (years)	Mean no. of interviews	% of men
SHARE	Austria	2004, 2006, 2011, 2013, 2015, 2017, 2019	2877	5.4	3.3	46.2
	Belgium	2004, 2006, 2011, 2013, 2015, 2017, 2019	4118	5.7	3.3	51.8
	Bulgaria	2017, 2019	377	2.0	2.0	43.0
	Croatia	2015, 2017, 2019	1119	2.8	2.4	49.5
	Cyprus	2017, 2019	124	2.0	2.0	42.7
	Czech Republic	2006, 2011, 2013, 2015, 2017, 2019	3827	5.8	3.4	41.4
	Denmark	2004, 2006, 2011, 2013, 2015, 2017, 2019	3031	6.6	3.5	48.2
ELSA	England	2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, 2018	9895	7.7	4.5	47.6
SHARE	Estonia	2011, 2013, 2015, 2017, 2019	3662	4.8	3.2	42.3
	Finland	2017, 2019	550	2.0	2.0	47.3
	France	2004, 2006, 2011, 2013, 2015, 2017, 2019	3540	6.4	3.4	47.1
	Germany	2004, 2006, 2011, 2013, 2015, 2017, 2019	3437	5.3	3.2	49.7
	Greece	2004, 2006, 2015, 2017, 2019	2187	6.5	2.7	60.5
	Hungary	2011, 2017, 2019	788	6.4	2.3	41.1
	Israel	2004, 2006, 2013, 2015, 2017, 2019	1447	8.2	3.3	46.9
	Italy	2004, 2006, 2011, 2013, 2015, 2017, 2019	3026	5.9	3.2	56.1
	Latvia	2017, 2019	303	2.0	2.0	41.6
	Lithuania	2017, 2019	528	2.0	2.0	37.1
	Luxembourg	2013, 2015, 2017, 2019	841	3.9	2.8	54.5
	Malta	2017, 2019	239	2.0	2.0	72.8
	Netherlands	2004, 2006, 2011, 2013, 2019	1862	6.4	2.7	55.6
	Poland	2006, 2011, 2015, 2017, 2019	1700	5.6	2.7	41.4
	Portugal	2011, 2015, 2017	761	4.6	2.3	50.5
	Romania	2017, 2019	560	2.0	2.0	45.5
	Slovakia	2017, 2019	665	2.0	2.0	47.7
	Slovenia	2011, 2013, 2015, 2017, 2019	2531	4.5	3.1	44.6
	Spain	2004, 2006, 2011, 2013, 2015, 2017, 2019	2731	5.3	3.0	59.6
	Sweden	2004, 2006, 2011, 2013, 2015, 2017, 2019	3151	6.2	3.2	45.1
	Switzerland	2004, 2006, 2011, 2013, 2015, 2017, 2019	2178	6.4	3.6	48.8
CRELES	Costa Rica	2005, 2007, 2009, 2010, 2012	1244	2.3	2.1	76.9
MHAS	Mexico	2001, 2003, 2012, 2015, 2018	8148	6.8	2.7	66.7
HRS	United States	1992, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008,	25 753	9.2	5.2	46.9

		2010, 2012, 2014, 2016, 2018				
CHARLS	China	2011, 2013, 2015, 2018	2819	4.8	2.9	54.1
JSTAR	Japan	2007, 2009, 2011	1775	3.0	2.5	64.8
KLoSA	South Korea	2006, 2008, 2010, 2012, 2014, 2016, 2018	5133	7.0	4.2	53.1
Overall			106 927	6.7	3.7	50.5

Table 1.2 presents the descriptive statistics by labor force status for 396,904 observations from 106,927 individuals, which consisted of 217,166 (54.7%) with working status and 179,738 (45.3%) with retired status.

Table 1.2 Descriptive Statistics of Observations by Labor Force Status

Variables, obs. (%)	Labor force status	
	Working (obs.=217 166)	Retired (obs.=179 738)
Age, years, mean (SD)	57.9 (4.7)	64.2 (4.3)
Men	113 377 (52.2)	83 958 (46.7)
Married	173 581 (79.9)	136 461 (75.9)
Education		
Low	50 673 (23.3)	53 159 (29.6)
Middle	93 304 (43.0)	83 463 (46.4)
High	55 364 (25.5)	35 754 (19.9)
Missing	17 825 (8.2)	7362 (4.1)
Job type		
Physical labor	104 359 (48.1)	39 952 (22.2)
Non-physical labor	77 364 (35.6)	37 235 (20.7)
Missing	35 443 (16.3)	102 551 (57.1)
Heart disease		
Ever had	16 158 (7.4)	31 688 (17.6)
Never	200 363 (92.3)	147 743 (82.2)
Missing or single observation	645 (0.3)	307 (0.2)
Stroke		
Ever had	3632 (1.7)	10 814 (6.0)
Never	212 884 (98.0)	168 607 (93.8)
Missing or single observation	650 (0.3)	317 (0.2)
Hypertension		
Ever had	73 349 (33.8)	89 924 (50.0)
Never	143 151 (65.9)	89 550 (49.8)
Missing or single observation	666 (0.3)	264 (0.2)
Diabetes		
Ever had	22 359 (10.3)	31 643 (17.6)
Never	194 099 (89.4)	147 756 (82.2)
Missing or single observation	708 (0.3)	339 (0.2)
Obesity		
BMI ≥30	44 184 (20.4)	44 339 (24.7)
BMI <30	147 527 (67.9)	115 958 (64.5)
Missing or single observation	25 455 (11.7)	19 441 (10.8)
Physical inactivity		
<1 per week	26 627 (12.3)	25 492 (14.2)
≥1 per week	116 137 (53.5)	104 568 (58.2)
Missing or single observation	74 402 (34.3)	49 678 (27.6)
Smoking		

Currently smoking	37 588 (17.3)	25 688 (14.3)
Not smoking	146 405 (67.4)	114 838 (63.9)
Missing or single observation	33 173 (15.3)	39 212 (21.8)
Binge drinking		
≥4/5 drinks per day	14 271 (6.6)	6449 (3.6)
<4/5 drinks per day	124 079 (57.1)	97 412 (54.2)
Missing or single observation	78 816 (36.3)	75 877 (42.2)

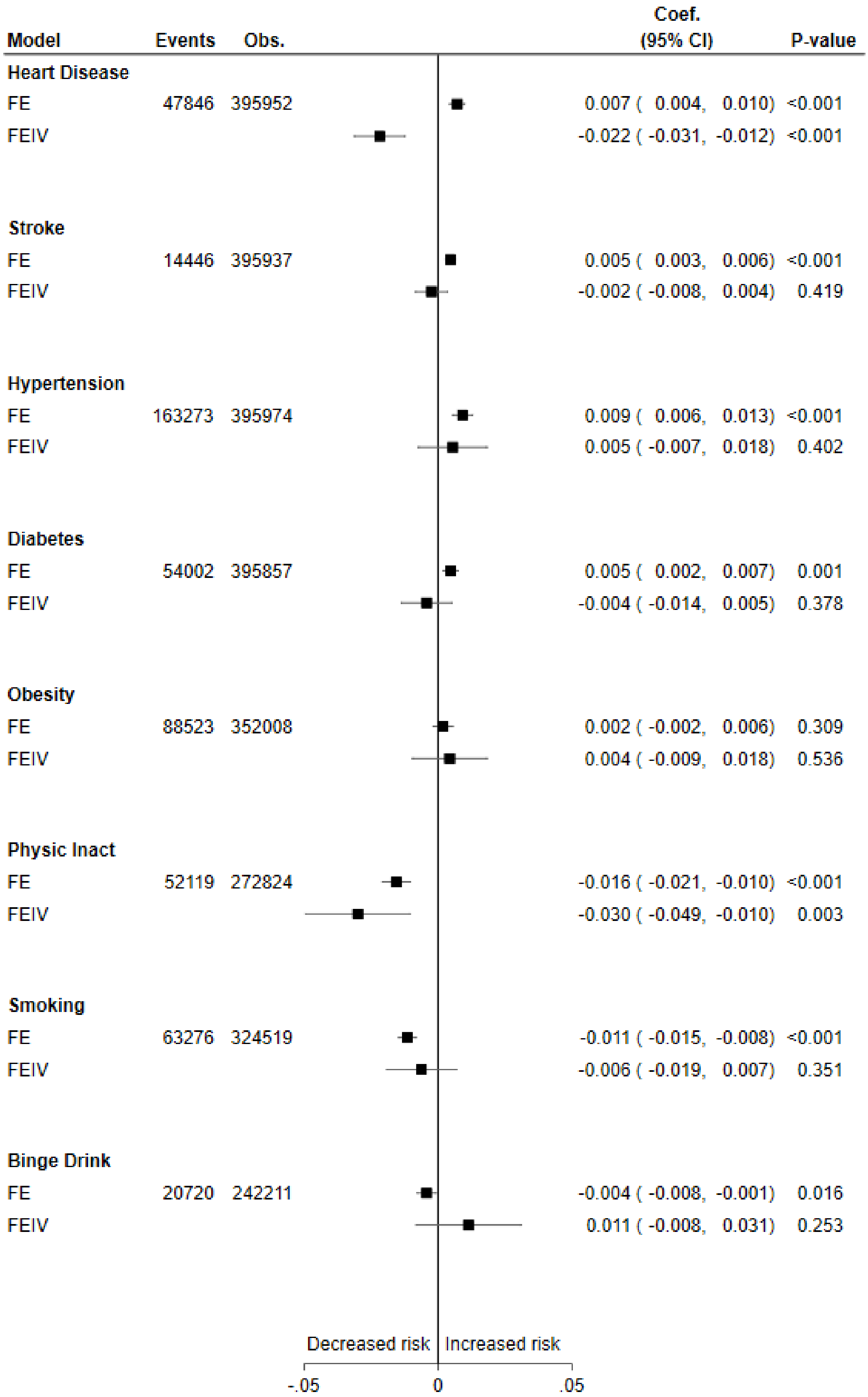
1.4.2 Estimation of FEIV Models

Before pooling data from different countries, we checked heterogeneity across countries. Country-by-country analyses using FEIV models indicated moderate heterogeneity in hypertension (I^2 statistic = 52.6%), diabetes (34.7%), and smoking (41.4%) (Appendix Figures F.1–8). After adjusting for country FEs, the test of interaction did not indicate signs of heterogeneity except for hypertension among Asian countries (Appendix Table F.2) and low-middle-income countries (Appendix Table F.3). Thus, it should be noted that the association between retirement and hypertension could be heterogeneous across countries, while other outcomes appeared to be homogeneous.

Figure 1.2 and Appendix Table F.4 present the results of the FE and FEIV models using pooled data. F statistics (Kleibergen & Paap, 2006) indicated that our IVs were strongly correlated with retirement, and the over-identification tests (Hansen, 1982) showed that the instruments of ERA and ORA were uncorrelated with residuals (Appendix Table F.4). In FE models without IVs, retirement was associated with an increased heart disease risk (coefficient = 0.007 [95% CI: 0.004 to 0.010]), stroke (0.005 [0.003 to 0.006]), hypertension (0.009 [0.006 to 0.013]), and diabetes (0.005 [0.002 to 0.007]). In contrast, the FEIV models showed a 2.2%-point decrease in the heart disease risk (-0.022 [-0.031 to -0.012]) as well as a 3.0%-point decrease in physical inactivity (-0.030 [-0.049 to -0.010]) among retirees, compared with workers. Readers concerned about multiple testing can interpret p-values using a Bonferroni correction (α = 0.05/8 outcomes = 0.006).

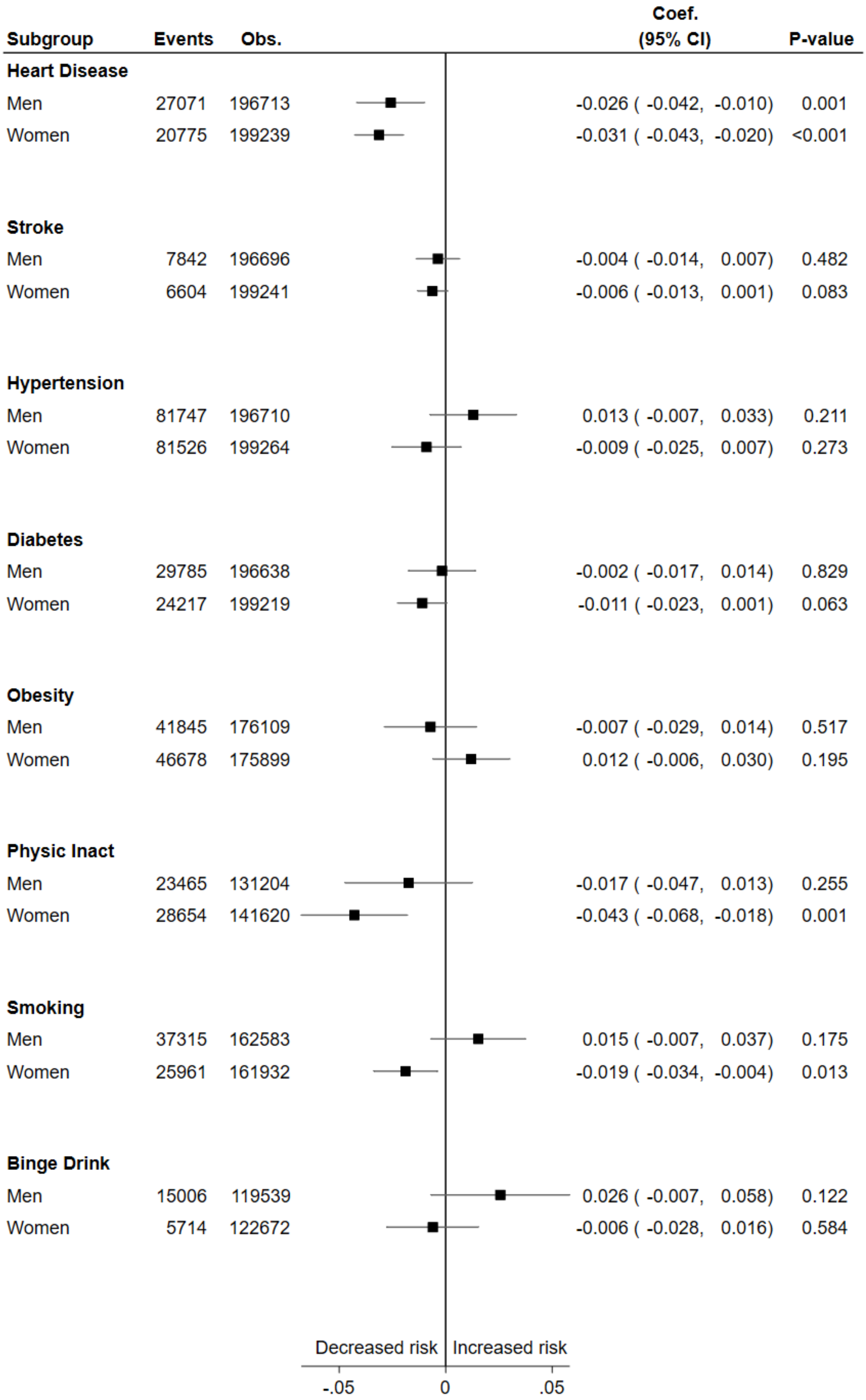
In subgroup analyses, we found some heterogeneous associations of retirement with CVD and risk factors. Figure 1.3 and Appendix Table F.5 present the results of subgroup analyses by sex using the FEIV. In both sexes, retirement was associated with a decreased heart disease risk. Among women, it was also associated with a 1.9%-point decrease in smoking (-0.019 [-0.034 to -0.004]). In Figure 1.4 and Appendix Table F.6, people with high educational levels showed associations between retirement and decreased risks of stroke (-0.014 [-0.028 to -0.001]), obesity (-0.029 [-0.057 to -0.001]), and physical inactivity (-0.045 [-0.080 to -0.011]). Figure 1.5 and Appendix Table F.7 present that people who retired from non-physical labor exhibited reduced risks of heart disease (-0.031 [-0.050 to -0.013]), obesity (-0.031 [-0.056 to -0.007]), and physical inactivity (-0.048 [-0.082 to -0.013]) compared with those continuing non-physical labor. In contrast, those retired from physical labor indicated an increased risk of obesity (0.025 [0.002 to 0.048]) compared with those engaging in physical labor.

Figure 1.2 Association of Retirement with Cardiovascular Diseases and Risk Factors



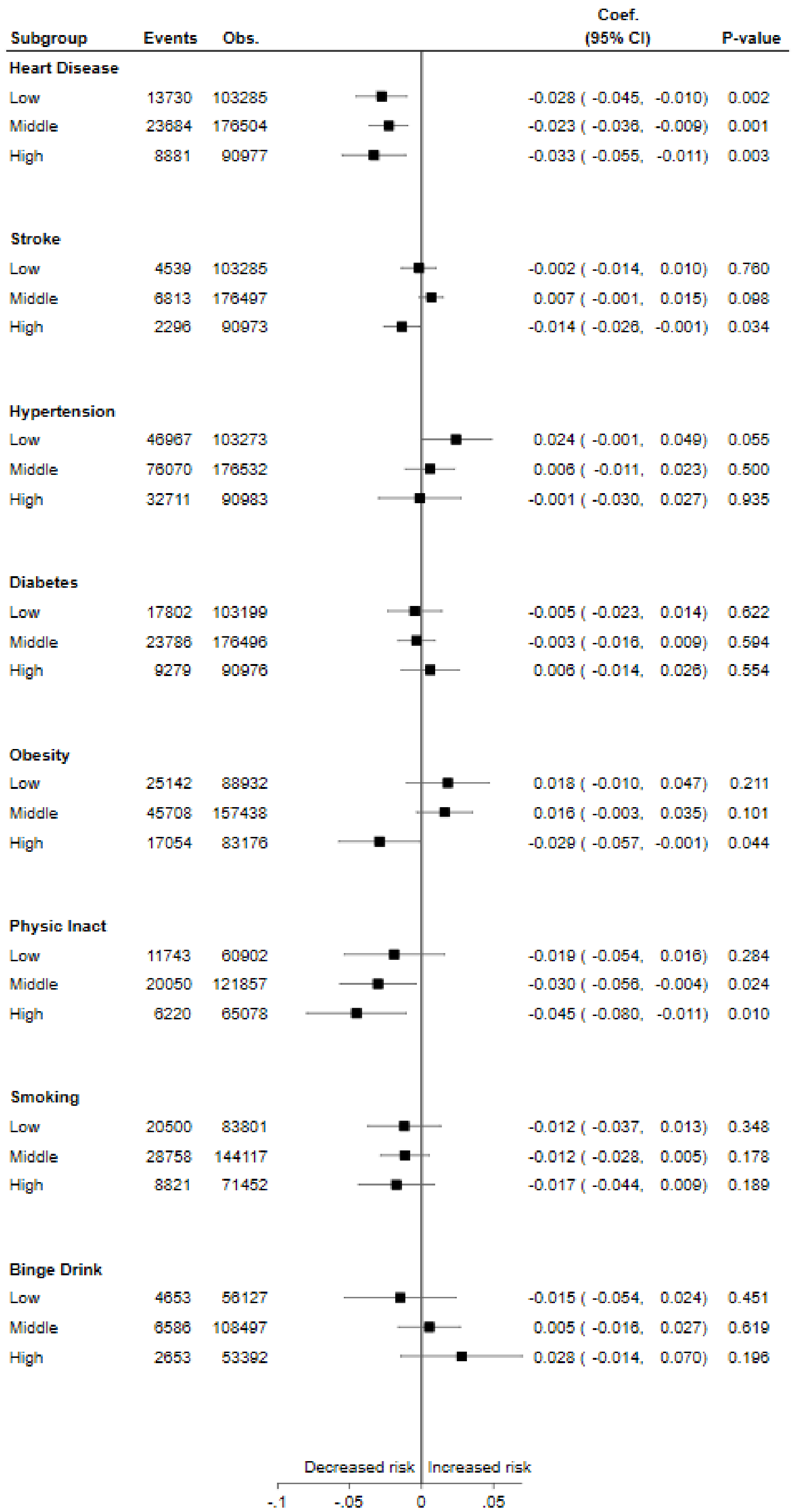
Physic Inact. denotes physical inactivity. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

Figure 1.3 Subgroup Analysis by Sex for the Association of Retirement with Cardiovascular Diseases and Risk Factors



Physic Inact. denotes physical inactivity. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

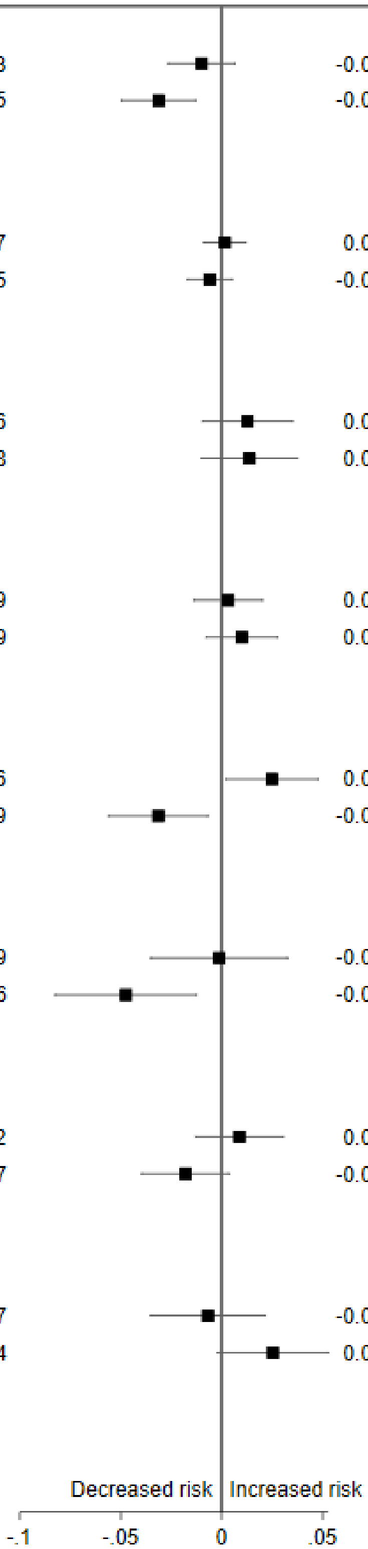
Figure 1.4 Subgroup Analysis by Education for the Association of Retirement with Cardiovascular Diseases and Risk Factors



Physic Inact. denotes physical inactivity. “Low” denotes less than upper secondary education; “Middle” denotes upper secondary and vocational training; “High” denotes tertiary education. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

Figure 1.5 Subgroup Analysis by Job Type for the Association of Retirement with Cardiovascular Diseases and Risk Factors

Subgroup	Events	Obs.	Coef. (95% CI)	P-value
Heart Disease				
Physical labour	14996	143993	-0.010 (-0.026, 0.007)	0.239
Non-physical labour	12295	114395	-0.031 (-0.050, -0.013)	0.001
Stroke				
Physical labour	3819	143987	0.002 (-0.009, 0.012)	0.765
Non-physical labour	3154	114385	-0.006 (-0.017, 0.005)	0.302
Hypertension				
Physical labour	54857	144006	0.013 (-0.009, 0.035)	0.258
Non-physical labour	44099	114398	0.014 (-0.010, 0.037)	0.254
Diabetes				
Physical labour	17133	143999	0.003 (-0.014, 0.020)	0.710
Non-physical labour	13388	114389	0.010 (-0.008, 0.028)	0.266
Obesity				
Physical labour	31522	129996	0.025 (0.002, 0.048)	0.032
Non-physical labour	25602	104209	-0.031 (-0.056, -0.007)	0.012
Physic Inact				
Physical labour	23459	110529	-0.001 (-0.035, 0.033)	0.942
Non-physical labour	12195	83696	-0.048 (-0.082, -0.013)	0.008
Smoking				
Physical labour	27086	125472	0.009 (-0.013, 0.031)	0.424
Non-physical labour	15234	98157	-0.018 (-0.040, 0.004)	0.105
Binge Drink				
Physical labour	10486	99077	-0.007 (-0.035, 0.022)	0.637
Non-physical labour	4880	75604	0.025 (-0.002, 0.053)	0.070



Physic Inact. denotes physical inactivity. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

1.4.3 Robustness Checks

First, we excluded the partly retired individuals and examined the association between full retirement and outcomes; the point estimates were similar to our main results though broader CIs were indicated in physical inactivity (Appendix Table F.8). Second, restricting the participants to those aged 52–68 years revealed similar associations of retirement with decreased heart disease risks and physical inactivity, while it showed amplified associations with decreased risks of diabetes, compared with the main results (Appendix Table F.9). Third, we excluded data from Greece, Latvia, Malta, Portugal, Romania, Costa Rica, Japan, and South Korea, in which the IVs appeared to be weak. The exclusion of these countries did not make a considerable difference to the results (Appendix Table F.10). Fourth, excluding data from the United States did not affect the results (Appendix Table F.11). Fifth, analysis using an imputed dataset showed similar results (Appendix Table F.12). Sixth, FEIV Poisson models indicated consistent results with linear probability models (heart disease: risk ratio = 0.89 [0.81 to 0.98]; physical inactivity: 0.87 [0.77 to 0.97]) (Appendix Table F.13). Seventh, we found consistent results even after excluding those who retired within two years (Appendix Table F.14). Eighth, the FEIV models including interactions between IVs and marital status in the first stage showed similar results (Appendix Table F.15). Finally, people who retired over five years ago exhibited larger reductions in heart disease risks and physical inactivity, whereas those who retired within five years presented a clearer reduction in stroke, compared with their counterparts (Appendix Figure F.9).

1.5 Discussion and Conclusions

In this multi-country longitudinal study, we examined the association of retirement with CVD and its associated risk factors using the IV method. The FE models without IVs showed an increased CVD risk among retirees and suggested that sicker workers retired earlier. In contrast, the FEIV models using the SPA as an IV showed the association of retirement with a decreased heart disease risk for the first time. These discrepancies in estimates between the FE and the FEIV models suggest the presence of the healthy worker survivor effect in previous research which showed the detrimental association of retirement with CVD risks (Xue et al., 2020). Our study provides updated results and highlights the importance of reconsidering the possible beneficial retirement effects on cardiovascular health. Our FEIV models also presented decreased physical inactivity after retirement, consistent with previous findings (Celidoni & Rebba, 2017; Eibich, 2015; Insler, 2014; Kesavayuth et al., 2018;

Müller & Shaikh, 2018; Syse et al., 2017). Physical activity may contribute to the decreased risk of heart disease among retirees.

The subgroup analyses revealed heterogeneous associations between retirement and risk factors. We found decreased smoking among women but not among men after retirement. Gender differences in workplace stress (Narayanan et al., 1999) and post-retirement social networks (Comi et al., 2022) might be correlated with cigarette consumption. Associations of retirement with reduced obesity and physical inactivity among people with high educational levels were shown. They may be aware of their health and afford to invest in health-promoting activities such as exercise and healthy eating after retirement. Although the analysis was limited to those who had working experience during the study period, people who retired from non-physical labor exhibited reduced obesity and physical inactivity, whereas those who retired from physical labor indicated increased obesity risk, in line with previous studies (Chung et al., 2009; Stenholm et al., 2017). Retirement provides time to undertake physical activity for those who have engaged in sedentary jobs, whereas those who have engaged in physical labor lose financial incentives for being physically active after retirement and may gain weight unless they continue exercising to the same extent as before. As a previous study suggested (Eibich, 2015), these differences in post-retirement health behaviors may be attributable to the differentiated CVD outcomes, specifically a reduced risk of stroke among highly educated people and heart disease among retirees from non-physical labor.

This study has several limitations. First, we could not determine the mechanism underlying the reduced CVD risk after retirement. Our analyses suggested that trends in CVD incidence and its risk factors were consistent; however, they were not conclusive because risk factor outcomes were only available in limited countries. Moreover, other important health behaviors (e.g., sleep duration, diet, and medication adherence) were also not available in the harmonized datasets. We assumed that hypertension, diabetes, and obesity are on the pathway between retirement and CVD but did not find clear patterns between retirement and these outcomes. Unobserved factors may have offset the beneficial effect of retirement on them though the net effect of retirement on heart disease appeared to be protective. Further studies are required to confirm the mechanism. Second, although harmonization across different surveys was performed by specialists in the field (Hu & Lee, 2012; Lee et al., 2021; Zamarro & Lee, 2012), some differences remained. Although these inconsistencies may induce bias in the estimates, some part of the potential bias could be reduced by adding the FEs of countries. Third, both retirement status and outcomes were self-reported and subject to measurement errors. Nevertheless, the direct question of retirement status is face-valid and reflects individuals' recognition of retirement. Such recognition plays an important role in adjusting post-retirement health behaviors (Eibich, 2015). Moreover, a previous study using data from the HRS showed that self-reporting of stroke has acceptable validity (Glymour & Avendano, 2009). If the reports of diagnosed health conditions were missed, it would tend to bias associations towards the null.

This novel study suggests that retirement was associated with a decreased heart disease risk on average. Some associations of retirement with CVD and risk factors appeared heterogeneous by individual characteristics. Policymakers need to consider the benefits of raising SPA and allowing older people to continue working versus the costs of the potential risk of expensive medical conditions such as CVD.

2. Sex Differences in the Impact of Retirement on Health: Evidence from 35 Countries[3]

2.1 Introduction

To accommodate an aging population, many developed countries are increasing their state pension age (SPA).[4] These policy changes may influence population health because they delay the timing of retirement and change budget constraints and time allocations for health investments in later life (Grossman, 1972). However, the impact of delayed retirement on health remains unclear, and there is a lack of consensus on this issue (Garrouste & Perdrix, 2022; Nishimura et al., 2018). We hypothesized that the inconsistent results in the previous literature stemmed from effect heterogeneity. When a subgroup adversely influenced by retirement conceals its beneficial effect in other subgroups, the average treatment effect of retirement in the population will be obscure. Therefore, this study examined the impact of retirement on health by employing an array of analyses targeting heterogeneity. This was accomplished through the utilization of harmonized longitudinal data derived from 35 countries.

Retirement is commonly an endogenous decision, as people in poorer health are more likely to retire. To address the potential downward bias, researchers often use SPA as an instrumental variable (IV) for retirement.[5] SPA seems to be valid because of its dual attributes; first, reaching the SPA will increase the probability of retirement (meeting the relevance condition); second, the SPA itself does not directly affect health outcomes (satisfying the exclusion restriction condition). Using the harmonized data, this study employed within- and between-country variations in SPA to identify the effects of retirement on health. Moreover, we benefitted from a panel structure of the longitudinal surveys and included fixed effects (FEs) of individuals, countries, years, and interactions between countries and years. Thus, our model, grounded in the fixed-effects instrumental variable (FEIV), effectively accounted for unobserved time-invariant characteristics of individuals and countries and heterogeneous time trends across countries. Additionally, this study delved into

---

[3] The results presented in this chapter have been presented as Sato & Noguchi (2023).
[4] See Organisation for Economic Co-operation and Development (2021).
[5] The regression discontinuity design (RDD) using SPA as a threshold for the timing of retirement is another possible identification strategy. Ebeid and Oguzoglu (2023) used the nonparametric fuzzy RDD and examined the effect of retirement on the cognitive function. Although the method has the strength of no parametric assumptions, we did not use it because it cannot account for multiple thresholds representing an early retirement age and the panel structure of longitudinal surveys (i.e., incompatible with fixed effects).

the disparities in the effects of retirement, aiming to unveil the underlying cause of the inconsistent results encountered in the earlier academic discourse.

We found a discernible pattern wherein women exhibited improved cognitive function and physical independence after retirement. In both sexes, retirement improved self-rated health, but this effect was more pronounced among women compared to men. Remarkably, this trend extended to a reduction in physical inactivity and smoking among women, which was not observed within the male cohort. Unlike the sex differences, the effects of retirement on the outcomes appeared homogeneous across various dimensions, including countries, educational backgrounds, and pre-retirement occupational characteristics.

This study contributes significantly to the existing literature in several ways. First, it suggests that the inconsistent results in previous studies can be attributable to sex heterogeneity in the effects of retirement. In particular, we found that retirement improved women's cognitive function, whereas men indicated an insignificant but detrimental effect of retirement. Intriguingly, in scenarios where a substantial proportion of male subjects constitutes the study population, their presence could potentially overshadow the beneficial consequences of retirement for women. This is consistent with many studies that showed no evidence of the association between retirement and cognitive function (Coe et al., 2012; Coe & Zamarro, 2011; Romero Starke et al., 2019; Rose, 2020). Our findings imply that the average treatment effect of retirement can vary depending on the sex composition prevalent within the study population.

Second, we present the potential mechanism of health disparities in the older population. The consistent sex differences in health outcomes and behaviors suggest that post-retirement health behaviors can induce the heterogeneous effects of retirement on health in line with Eibich (2015). We showed that retirement is beneficial, especially for women; thus, delayed retirement owing to increasing SPA can deteriorate population health. However, increasing SPA seems inevitable in many developed countries given their imminent pension finance facing the challenges posed by the rapidly aging population. Our findings provide policymakers with valuable insights that the promotion of healthy behaviors can mitigate potential adverse effects of delayed retirement on health owing to the mounting SPA.

Third, we provide an encompassing perspective on the effects of retirement vis-à-vis health. A recent review showed that many studies indicated the detrimental effects of retirement on cognitive function, while evidentiary support for its influence on physical function remains inconclusive (Garrouste & Perdrix, 2022). Notably, retirement appears to yield beneficial consequences for self-rated health (Garrouste & Perdrix, 2022).

Paradoxically, this array of findings across diverse outcomes remains enigmatic, as self-rated health constitutes a strong predictor of both cognitive and physical impairments (Bond et al. 2006; Brenowitz et al. 2014; Idler and Benyamini 1997). Inconsistencies in prior literature may be owing to variations in statistical methodologies, study designs, retirement measurement modalities, outcome assessments, and contexts of various countries under examination. The discrepancy lacks a clear explanation, and thus a comprehensive cross-country investigation using consistent research methodologies is required. One exception is Nishimura, Oikawa, and Motegi (2018), who explored the effects of retirement on various health outcomes using data from eight countries up to 2014.[6] Our study expands the work of Nishimura, Oikawa, and Motegi (2018) in several ways. Foremost, they investigated each country separately, leaving unsolved the issue of cross-country variations in the health implications of retirement. Contrarily, we tested the effect heterogeneity by countries by including interaction terms between retirement and the characteristics of country, such as region, income level, and the percentage of the older population. We confirmed that the effect of retirement was homogeneous across countries and thus showed pooled estimates from data of 35 countries. Moreover, given that many countries started raising their SPA around 2015 (Organisation for Economic Co-operation and Development, 2021), this study applies more recent data from a larger number of countries. Finally, we explored not only health outcomes but also health behaviors including physical inactivity, smoking, and binge drinking, which enabled us to reveal the underlying mechanism that contributes to heterogeneous effects on health outcomes.

The remainder of this chapter is organized as follows: Section 2.2 outlines the data used in this study; Section 2.3 presents the empirical model; Section 2.4 reports the results; and Section 2.5 discusses the results and concludes the paper.

2.2 Data

2.2.1 Harmonized Data

We used the same datasets as described in Chapter 1. We implemented the harmonized datasets of the Health and Retirement Study (HRS) and its sister surveys provided by the Gateway to Global Aging Data project (Lee, Phillips, and Wilkens 2021), which is "a free public resource designed to facilitate cross-national and longitudinal studies on aging." It includes data on various individual characteristics regarding demographics, health, health

[6] The aforementioned data source pertains to the HRS, the SHARE, the ELSA, the JSTAR, and the KLoSA.

services use, work/employment, economic status, and family structure/social network.[7] Data from the Longitudinal Aging Study in India and the Malaysia Ageing and Retirement Study were not used because they were harmonized only for one wave. Data from the Irish Longitudinal Study on Ageing were also excluded because harmonized variables related to activities of daily living (ADL), instrumental activities of daily living (IADL), and binge drinking were provided only for wave 1. Given this, our datasets comprised of waves 1, 2, and 4 through 8 (2004–2019) of the Survey of Health, Ageing and Retirement in Europe (SHARE)[8]; waves 1 through 9 (2002–2018) of the English Longitudinal Study on Ageing (ELSA); waves 1 through 5 (2005–2012) of the Costa Rican Longevity and Healthy Aging Study (CRELES); waves 1 through 5 (2001–2018) of the Mexican Health and Aging Study (MHAS); waves 1 through 14 (1992–2018) of the HRS; waves 1 through 4 (2011–2018) of the China Health and Retirement Longitudinal Study (CHARLS); waves 1 through 3 (2007–2011) of the Japanese Study of Aging and Retirement (JSTAR); and waves 1 through 7 (2006–2018) of the Korean Longitudinal Study of Aging (KLoSA). Finally, the CRELES included a cohort interviewed in waves 1 through 3 (2005–2009) and another cohort interviewed in waves 4 through 5 (2010–2012). The surveys were all designed to represent the national older population except for the JSTAR, which randomly recruited participants from 10 specific municipalities. The same individuals were interviewed biennially; however, the MHAS and CHARLS conducted interviews triennially since 2012 and 2015, respectively.

Originally, the harmonized data involved 1,912,071 observations from 276,930 unique individuals who were surveyed in multiple timings. Then, our inclusion and exclusion criteria were applied. First, we included 609,422 observations from 205,022 unique individuals aged 50 to 70 years, whose timings could be affected by the SPA of each country.[9] Second, regarding the CHARLS, 39,682 observations from rural residents were not included because China had different pension systems in rural and urban areas; therefore, rural residents were not affected by the SPA (Lei and Liu 2018). Third, we excluded 118,100 observations that corresponded to individuals who were not working for reasons other than retirement, such as being unemployed, disabled, or a homemaker. Fourth, we did not include

---

[7] GATEWAY TO GLOBAL AGING DATA. https://g2aging.org/ (Accessed: January 21, 2023)

[8] SHARE was conducted in 29 countries, namely Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, England, Estonia, Finland, France, Germany, Greece, Hungary, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and Switzerland. The third wave of the SHARE, referred to as SHARELIFE, featured a distinct questionnaire from the preceding waves.

[9] The CRELES conducted interviews with individuals aged ≥ 60 years in waves 1 through 3, and with individuals aged 55 through 65 years in waves 4 and 5. Whereas the HRS interviewed adults aged ≥ 51 years.

355 observations because of missing values for explanatory variables necessary for our analysis. Finally, 54,381 individuals who were observed only once in the survey were excluded from analyses because maintaining them in a FE model could underestimate standard errors (Correia 2015). Thus, at the baseline, our study consisted of 396,904 observations from 106,927 individuals in 35 countries with a mean follow-up period of 6.7 years (Table 1.1). Notably, the number of observations varied across regressions, as those with missing values, differed across outcomes.

2.2.2 Retirement and State Pension Age

The retirement statuses of survey participants were determined using the harmonized variable of self-reported labor force status in Appendix C and described by Zamarro and Lee (2012). Individuals who self-identified as retired during the interview, regardless of their working status (i.e., including those who were "partly retired"), were included in the retired group for comparison with workers, as defined in previous literature (Atalay et al., 2019; Bianchini & Borella, 2016), and outlined in Appendix C. Other studies defined retirement as not working (Bingley & Martinello, 2013; Bonsang et al., 2012; Coe & Zamarro, 2011). In light of this alternative definition, individuals who identified as retired but were still engaged in paid work were classified by combining labor force status and employment engagement variables, and subsequently excluded from the analysis in a robustness check.

To address potential endogeneity in retirement decisions, we employed the SPA as an IV for retirement. In some countries, early pensions are granted under specific circumstances, such as reduced benefits or sufficient social security contributions. Thus, we employed the joint instruments of the early retirement age (ERA) and the official retirement age (ORA) to predict retirement. A binary ERA variable indicated whether participants had attained the earliest age of eligibility for reduced pensions or full pensions with certain conditions. Similarly, a binary ORA variable indicated whether participants had attained the age of entitlement to minimum guaranteed pensions or full pensions without any requirements. In countries where early pensions are not available, the ERA variable was set to zero for all participants. We obtained ERA, ORA, and their modifications during the study period from "Social Security Programs Throughout the World" (United States Social Security Administration, 2020), "Pensions at a Glance" (Organisation for Economic Co-operation and Development, 2021), and websites of the national authorities (Appendix D). To provide a descriptive overview of retirement patterns, we created a graph depicting age and the corresponding retirement rate by country, which demonstrated changes in the retirement rate around the SPA (Appendix E).

2.2.3 Outcome Measures

As a measure of cognitive function, we focused on episodic memory involving a neurocognitive system that is responsible for recollecting past experiences. Episodic memory constitutes an appropriate measure to assess the impact of retirement because it exhibits a decline with advancing age (Tulving 2002) and can capture the preliminary stages of cognitive impairments. Moreover, it is less subjective to floor and ceiling effects, given a wide range of scores (Bonsang et al. 2012). In accordance with the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) battery (Morris et al., 1989), a list of common words was verbally presented to the participants, following which they were immediately asked to recollect as many words from the list as possible. After approximately 5 min, the participants were requested to, once again, recall the words from the list. The episodic memory score was calculated by adding the number of words remembered during both the immediate and delayed recalls. Typically, most surveys included a list of ten words, thereby offering a score range from 0 to 20. As shown in Appendix Figure G.1, the scores appeared to be normally distributed for surveys featuring a 10-word list. However, the number of words varied across surveys and waves. Waves 1 through 2 of the HRS comprised 20 words on the list, while the MHAS consisted of 8 words, and the CRELES and the KLoSA contained 3 words (Shih, Lee, and Das 2012). Thus, to enable comparison, we standardized the scores to z-scores[10] for each survey. Additional analysis, using only surveys with a 10-word list, was performed for robustness.

The assessment of physical function in our study was based on the individual's capability to carry out ADL and IADL. A combination of ADL and IADL items is known to accurately predict physical limitations (Roehrig, Hoeffkin, and Pientka 2007). To facilitate comparability, we selected eight activities that participants were capable of performing. These activities included four ADL items (bathing, eating, getting in and out of bed, and using the toilet) and four IADL items (managing money,[11] taking medications, shopping for groceries,

---

[10] The z-score is a score converted so that the mean is 0 and the standard deviation is 1, which makes values with different units of measurement (such as outcome measures in the harmonized data) comparable.

[11] The JSTAR contained three distinct queries concerning financial management, namely paying bills, withdrawing money from the bank, and filling out a pension document. To ensure consistency with other surveys, we used the first inquiry to create the variable for the Japanese participants.

and preparing meals). As wave 1 of the HRS did not include questions about the capability of toilet use and four IADL items, the score was recorded as missing. The responses to the eight items displayed consistency (Cronbach's α = 0.79). As shown in Appendix Figure G.2, most participants were capable of performing all eight activities. Hence, we categorized the participants into two groups: those who were independent in all activities and those who were not. Subsequently, we created a binary variable indicating 1 for those who were fully independent and 0 for otherwise.

Self-rated health was measured using a 5-point Likert scale (1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent). In some prior studies, self-rated health was dichotomized into a binary variable indicating good or poor health (Behncke 2012; Coe and Lindeboom 2008; Hessel 2016; Johnston and Lee 2009; Messe and Wolff 2019; Neuman 2008; Rose 2020; Zhu 2016). However, as it appeared normally distributed (as illustrated in Appendix Figure G.3), this study standardized self-rated health to z-scores for each survey and treated it as a continuous variable, similar to other studies (Calvo, Sarkisian, and Tamborini 2013; Gorry, Gorry, and Slavov 2018).

Finally, to explore the underlying mechanisms linking retirement to the primary outcomes, we also examined the impact of retirement on physical inactivity, smoking, and binge drinking as these factors have been identified as potential risk factors for cognitive and physical impairments (Agahi et al. 2018; Maurage et al. 2012; Moore, Endo, and Carter 2003; Okusaga et al. 2013; Sato et al. 2021). Those who engaged in vigorous or moderate physical activity less than once per week were considered physically inactive individuals.[12] The smoking status of participants was categorized into current smokers and non-smokers.[13] Binge drinking was defined as consuming five or more drinks per day for men and four or

[12] In the case of South Korea, we relied solely on the variable of pertaining to vigorous physical activity since the KLoSA did not inquire about the frequency of moderate physical activity. For wave 7 of SHARE, only individuals who had also participated in wave 3 were asked about the frequency of physical activity. Thus, all observations from Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania, and Slovakia were excluded from the analysis because those only had one observation. In waves 1 through 3 of CHARLS, only half of the participants were queried about physical activity, and certain observations were excluded owing to question incompatibility. Specifically, the MHAS asked whether participant engaged in vigorous physical activity three or more times per week; in contrast, the JSTAR asked for the number of minutes of exercise on weekdays and weekends.

[13] In wave 6 of SHARE, individuals who had previously been interviewed were not queried about their smoking status. In wave 7, only new participants and those who previously reported smoking in wave 3 were asked about their current smoking status. Consequently, all the observations from Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania, and Slovakia were excluded from the analysis owing to a single observation.

more for women, in accordance with the definition provided by Centers for Disease Control and Prevention (2022).[14] These three variables were converted into binary categories.

2.2.4 Covariates

The estimation model utilized in this study incorporated adjustments for covariates including age, age squared (divided by 10), and marital status. Age was centered at the mean in regression models for ease of interpretation. During each interview, participants were asked to report their marital status, with the response options being: married, partnered, separated, divorced, widowed, or never married. We coded 1 for those who were married or partnered and 0 otherwise.

To access potential effect heterogeneity across various demographic and occupational characteristics, we included interaction terms between retirement and sex, educational levels, pre-retirement job characteristics, and country characteristics in our statistical models. Educational attainment was classified into three groups using the 1997 International Standard Classification of Education codes—less than upper secondary education as low, upper secondary and vocational training as middle, and tertiary education as high. We also investigated whether retirement from physically demanding jobs and jobs with low control modified the impact of retirement on the outcomes. During the surveys, participants who were currently employed were asked to rate their agreement with the following statements regarding the physical demands[15] and control of their job[16]; “My job is physically demanding” and “I have very little freedom to decide how I do my work.” Each item was measured using a four-point Likert scale that included “strongly disagree,” “disagree,” “agree,” or “strongly agree.” Participants who responded with “agree” or “strongly agree” at least once during the interview were considered to have experience in physical labor and low-control jobs, respectively. Participants who had never engaged in paid work during the study period were excluded from the models with job characteristics interactions. Additionally, we performed interaction tests across country characteristics, grouping regions into Europe (including Israel), America (Costa Rica, Mexico, and the United States), and Asia (China,

[14] Waves 1 and 6–8 of SHARE, wave 1 of ELSA, waves 1 and 2 of HRS, and all waves of CRELES and CHARLS did not provide the information on the number of drinks per day. Thus, all observations from Bulgaria, Croatia, Cyprus, Finland, Greece, Hungary, Latvia, Lithuania, Luxembourg, Malta, Portugal, Romania, and Slovakia were excluded from the analysis since these individuals had only one observation. Additionally, the question on drinking habits was not asked in wave 2 in some study sites of JSATR.

[15] The question was not included in wave 1 of the ELSA, CRELES, MHAS, and CHARLS.

[16] The question was not included in wave 1 of the ELSA; CRELES; MHAS; and waves 1 through 7 of HRS, CHARLS, JSTAR, and KLoSA.

Japan, and South Korea); classifying countries as high-income (all European countries, United States, Japan, and Korea) or low-middle income (Bulgaria, Romania, Costa Rica, Mexico, and China) based on Gross National Income per capita as defined by the World Bank[17]; and considering a country to be an aged society if the percentage of the population aged 65 years and older[18] exceeded 14%, as defined by the Organisation for Economic Co-operation and Development and World Health Organization (2020).

2.3 Empirical Model

We investigated the impact of retirement on the outcomes using linear probability models estimated by the FEIV with the two-stage least squares procedure (we fitted the same model shown in Section 1.3.1, so the equations were omitted). In all analyses, we estimated robust standard error clustering for individual, country, year, and interactions between country and year.

The FEIV model has several advantages with respect to pooling data from different countries and estimating the causal effect of retirement on outcomes. The FEs of individuals and countries provide controls for both observable and unobservable time-invariant factors, such as genetic predisposition and educational attainment, as well as institutional and cultural differences among countries. Additionally, the interaction between FEs and countries represents heterogeneous time trends across countries. Moreover, we applied the IV method using the ERA and ORA as instruments to mitigate the endogeneity of retirement. To be valid, IVs must meet two conditions, namely (i) the relevance condition (the IV is associated with treatment, i.e., retirement) and (ii) the exclusion restriction (the IV has no association with potential outcomes under different values of treatment). Furthermore, the assumption of monotonicity (i.e., the IV does not have conflicting effects on treatment in any individual) enables us to interpret the point estimate of $\hat{\gamma}_1$ as a local average treatment effect (LATE) among "compliers" (i.e., individuals who would retire upon reaching the SPA).

---

[17] Although Poland did not meet the threshold for high-income countries in 2006, we categorized it as a high-income country through the study period.
[18] We obtained the percentage of the population aged 65 and older from "World Development Indicators" published by the World Bank (2022).

2.4 Results

2.4.1 Descriptive Statistics

Table 2.1 presents the descriptive statistics by labor force status for a large sample of 396,904 observations from 106,927 individuals, which consisted of 217,166 (54.7%) with working status and 179,738 (45.3%) individuals with a retired status. At first glance, Table 2.1 indicates that retirement seems to deteriorate all health statuses, although it would be beneficial for health behaviors as risk factors. Notably, retirees were found to be older and less likely to be men, married, and highly educated. Finally, retirees are less likely to have experienced physical labor and a job with low control than workers.

Table 2.1 Descriptive Statistics of Observations by Labor Force Status

Variables	Working (obs.=217,166)			Retired (obs.=179,738)			Difference
	Obs.	Mean	SD	Obs.	Mean	SD	
Outcome variables							
Health status							
Cognitive function (z-score)	204,541	0.172	0.952	172,735	-0.0595	0.990	0.232***
Physical independence	194,608	0.959	0.198	168,365	0.883	0.322	0.076***
Self-rated health (z-score)	209,867	0.263	0.913	174,764	-0.105	0.986	0.368***
Health behavior as risk factors							
Physical inactivity	142,764	0.187	0.390	130,060	0.196	0.397	-0.009***
Smoking	183,993	0.204	0.403	140,526	0.183	0.387	0.021***
Binge drinking	138,350	0.103	0.304	103,861	0.062	0.241	0.041***
Covariates							
Age	217,166	57.940	4.694	179,738	64.217	4.253	-6.278***
Married	217,166	0.799	0.401	179,738	0.759	0.428	0.040***
Potential effect of heterogeneity							
Men	217,166	0.522	0.500	179,738	0.467	0.499	0.055***
Education							
Low	199,341	0.254	0.435	172,376	0.308	0.462	-0.054***
Middle	199,341	0.468	0.499	172,376	0.484	0.500	-0.016***
High	199,341	0.278	0.448	172,376	0.207	0.405	0.071***
Physically demanding job	181,723	0.574	0.494	77,187	0.518	0.500	0.056***
Low control job	121,760	0.355	0.479	44,475	0.323	0.468	0.032***

Unpaired t-tests were performed. *** p<0.01, ** p<0.05, * p<0.1.

As previously mentioned, we only included individuals who were followed up with at least twice, which may introduce selection bias caused by attrition. To assess this potential impact, we compared characteristics between individuals who were followed up and those who were lost in the follow-up (Appendix Table G.1). Consequently, we confirmed almost no differences in characteristics between these two groups; however, we observed that those who were lost in the follow-up were on average 0.84 years older and had poorer self-rated health by 0.11 points than those who were followed-up. Thus, exercising caution is necessary to interpret our results of age and subjective health.

2.4.2 Overall Regression Results

Tables 2.2 and 2.3 present comprehensive analyses of the effects of retirement on health outcomes and health behaviors as risk factors, respectively. Regressions were adjusted for age, age squared, and marital status. In Table 2.2, the point estimates obtained from pooled ordinary least squares (OLS) and FE models demonstrate that retirement has negative effects on health outcomes. However, the FEIV estimates reveal positive effects on health outcomes, with the difference between these models being the treatment of the decision to retire as an endogenous variable. These findings suggest that the inconsistencies in the effects of retirement on health outcomes in prior studies may partially be attributed to variations in estimation methods. Moreover, the FEIV estimates in our study do not show the paradoxical evidence observed in previous investigations, which indicate that retirement enhances self-rated health but has detrimental impacts on cognitive and physical functions. Specifically, our results indicate that retirement significantly improves the z-score of cognitive function by 0.048 standard deviations (SD), the likelihood of physical independence by 2.7% points, and the z-score of self-rated health by 0.144 SD. To elucidate the mechanisms underlying these results, we further evaluate the effect of retirement on health behaviors as risk factors. Table 2.3 shows that retirees exhibit a 3.0%-point reduction in physical inactivity compared with workers in the FEIV model. In the pooled OLS model, retirement was associated with increased smoking, which could be confounded with various factors such as education and mental health.[19] The association flipped to be negative after adjusting for FEs and finally became insignificant in the FEIV model. The pooled OLS model indicated the association of retirement with decreased binge drinking, whereas the FEIV model did not show a significant effect.

For all FEIV models, the Kleibergen-Paap Wald F statistics (Kleibergen and Paap 2006) indicated a strong correlation between IVs and retirement. In addition, the over-identification tests (Hansen 1982) did not reject the null hypotheses at the 5% significance level that the instruments were uncorrelated with residuals, which indicates that the IVs are plausible and satisfy the requirements for being valid. The first stage estimates of FEIV

[19] People with low education are more likely to retire as shown in Table 2.1 and to smoke than those with high education (Andersen et al., 2023). Similarly, people with mental illness are more likely to retire and smoke than those without mental illness (Steinberg et al., 2015).

models are presented in Appendix Table G.2. It shows that reaching ERA and ORA had significantly positive effects on the probability of retirement, which suggests that raising the ERA or ORA would delay the timing of retirement. The effect of ORA was more pronounced than that of ERA. The probability of retirement increased with age; conversely, the negative interaction terms of ERA and ORA with age indicate that the slope slowed down once ERA and ORA were reached.

Table 2.2 Associations Between Retirement and Health Outcomes

	Cognitive function			Physical independence			Self-rated health		
	Pooled OLS	FE	FEIV	Pooled OLS	FE	FEIV	Pooled OLS	FE	FEIV
Retirement	-0.104***	-0.002	0.048**	-0.087***	-0.023***	0.027***	-0.326***	-0.055***	0.144***
	(0.022)	(0.005)	(0.021)	(0.011)	(0.003)	(0.006)	(0.021)	(0.008)	(0.019)
Individual FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Country FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Year FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Country x Year FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
IV	NO	NO	YES	NO	NO	YES	NO	NO	YES
Observations	377,276	377,276	377,276	362,973	362,973	362,973	384,631	384,631	384,631
Adjusted R^2	0.023	0.478		0.027	0.418		0.041	0.564	
Kleibergen-Paap F			2230.315			2222.211			2184.288
Hansen J			0.684			0.313			2.510

All regressions are adjusted for age, age squared, and marital status. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.3 Associations Between Retirement and Health Behaviors

	Physical inactivity			Smoking			Binge drinking		
	Pooled OLS	FE	FEIV	Pooled OLS	FE	FEIV	Pooled OLS	FE	FEIV
Retirement	-0.004	-0.016***	-0.030***	0.020**	-0.011***	-0.006	-0.026***	-0.004**	0.011
	(0.024)	(0.005)	(0.010)	(0.008)	(0.002)	(0.007)	(0.009)	(0.002)	(0.010)
Individual FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Country FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Year FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
Country x Year FE	NO	YES	YES	NO	YES	YES	NO	YES	YES
IV	NO	NO	YES	NO	NO	YES	NO	NO	YES
Observations	272,824	272,824	272,824	324,519	324,519	324,519	242,211	242,211	242,211
Adjusted R^2	0.002	0.418		0.012	0.762		0.008	0.448	
Kleibergen-Paap F			1855.314			1592.798			788.561
Hansen J			2.471			3.627*			0.640

All regressions are adjusted for age, age squared, and marital status. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

2.4.3 FEIV Models Incorporating Interactions

To determine the extent of heterogeneity, interaction terms between retirement and several demographic, socio-economic, and other contextual factors were included in the FEIV, such as sex (Appendix Table G.3), educational levels (Appendix Table G.4), pre-retirement job characteristics including physically demanding job (Appendix Table G.5) and a job with low control (Appendix Table G.6), region (Appendix Table G.7), country income levels (Appendix Table G.8), and population aging rates (Appendix Table G.9). In summary, most of the interaction terms failed to achieve statistical significance, indicating that retirement has a homogenous impact on health outcomes and behaviors across these characteristics. An encouraging development is that the homogenous outcomes observed across diverse country attributes, including economic status and demographic trends related to aging, added support to the authenticity of our investigation, which entailed aggregating data from numerous countries. However, there was a heterogeneous effect on cognitive function, self-rated health, and smoking across sexes, as demonstrated in Appendix Table G.3. Compared with women, retirement was found to be less likely to enhance the z-score of cognitive function by 0.088 SD and that of self-rated health by 0.071 SD, while it was more likely to increase smoking by a 4.8% point among men. Nevertheless, it should be noted that the over-identification test did not meet the requirement for self-rated health and physical inactivity at the 5% significance level. Additionally, Appendix Table G.6 suggests that the characteristics of pre-retirement job control modified the effect of retirement on smoking.

2.4.4 Stratified Analysis by Sex

Given the notable interactions between retirement and sex in the preceding section, we performed a stratified analysis based on sex. The results are presented in Table 2.4. Among men, we found no significant association, except for self-rated health; male retirees demonstrated a 0.100 SD increase in self-rated health. Conversely, female retirees showed a 0.100 SD increase in cognitive function, a 3.8%-point increase in physical independence, and a 0.193 SD increase in self-rated health with respect to health outcomes. Moreover, female retirees curtailed their physical inactivity by 4.3% points and smoking by 1.9% points with respect to health behaviors. Considering the potential effect modification for the association between retirement and smoking, we also performed stratified analysis by sex and job control. Table 2.5 shows that women who retired from high-control jobs drove the effect of retirement on reduced smoking. Based on the results, the mechanism of retirement to enhance health behavior and, consequently, health outcomes, would be more lucid for women than men.

Table 2.4 Stratified FEIV Models for the Effect of Retirement on Health Outcomes and Behaviors by Sex

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
			Men			
Retirement	-0.005	0.015	0.100***	-0.017	0.015	0.026
	(0.032)	(0.009)	(0.030)	(0.015)	(0.011)	(0.017)
Observations	183,386	175,722	190,480	131,204	162,583	119,539
Kleibergen-Paap F	925.970	912.002	911.613	745.062	643.583	335.596
Hansen J	1.568	2.927*	1.006	0.115	2.118	0.098
			Women			
Retirement	0.100***	0.038***	0.193***	-0.043***	-0.019**	-0.006
	(0.027)	(0.008)	(0.024)	(0.013)	(0.008)	(0.011)
Observations	193,890	187,251	194,151	141,620	161,932	122,672
Kleibergen-Paap F	1333.937	1344.962	1300.021	1142.730	977.660	481.215
Hansen J	0.000	0.493	2.049	4.886**	0.002	0.711

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.5 Stratified FEIV Models for the Effect of Retirement on Smoking by Sex and Job Control

	Men & Low-control	Men & High-control	Women & Low-control	Women & High-control
Retirement	0.018	-0.003	0.021	-0.026*
	(0.025)	(0.020)	(0.018)	(0.015)
Observations	20,918	42,242	23,214	45,127
Kleibergen-Paap F	113.032	123.423	170.626	209.736
Hansen J	0.001	1.067	0.737	0.000

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

2.4.5 Robustness Check

We performed additional analyses to ascertain the robustness of our findings. First, we excluded participants who claimed to be retired but were still working, or partly retired, thereby considering the alternative definition of retirement i.e., fully retired. Although the point estimates were similar to the main results, the association between retirement and cognitive function was not statistically significant (Appendix Table G.10).[20]

Second, given that many participants were retired throughout the study period, we restricted the sample to those who reported being engaged in paid work at least once in the interviews (Model 1 in Appendix Table G.11). From Model 1, we excluded those who were self-employed (as shown in Model 2) and further eliminated those who engaged in a part-time job[21] (as demonstrated in Model 3). The results were almost consistent with the main findings even though the associations of retirement with cognitive function and physical inactivity were no longer statistically significant in Model 3.

Third, we narrowed the age range of participants between 52 and 68 years old. While the association between retirement and cognitive function became statistically insignificant, other outcomes were consistent with the main result (Appendix Table G.12).

Fourth, we present the outcomes of country-by-country analyses in Appendix Tables G.13–18, which revealed that certain countries, namely Greece, Latvia, Malta, Portugal, Romania, Costa Rica, Japan, and South Korea, had weak IVs with F statistics below the Stock-Yogo's critical value of 10% maximal relative bias (Stock and Yogo 2002). These were excluded from the analysis. The results obtained after exclusions (Appendix Table G.19) were similar to the main findings. Moreover, given that 24.1% of the participants were from the United States, we excluded data from the HRS and confirmed through Appendix Table G.20 that this exclusion did not affect the outcomes. In addition, countries with no changes in their SPA during the study period, namely Cyprus, Finland, Luxembourg, Sweden, Switzerland, Costa Rica, Mexico, China, and Japan, were excluded from the analysis as they did not have a within-country variation in FEIV estimation, and the results obtained were consistent with the main findings (Appendix Table G.21).

Fifth, as mentioned above, the length of a word list in the cognitive function test differed by surveys, and thus, we restricted our analysis to surveys with a 10-word list and investigated the association between retirement and the raw scores of cognitive function

[20] Nishimura, Oikawa, and Motegi (2018) also demonstrated that the robustness of retirement effects estimates across different definitions of retirement in their replication of previous studies.

[21] We defined part-time work as labor force participation of fewer than 35 hours per week.

scores. The results revealed no clear association among men, while female retirees could recall 0.281 more words than workers, which is in line with the main findings (Appendix Table G.22).

Sixth, to mitigate potential bias from missing observations, we adopted multiple imputations using the algorithm of expectation–maximization with bootstrapping (Honaker and King 2010) and created ten imputed datasets.[22] Although the association between retirement and cognitive function was not statistically significant using imputed data, other outcomes were similar to the main results (Appendix Table G.23).

Seventh, we investigated short-term and long-term retirement effects, as previous studies have suggested that these effects are dynamic (Blake and Garrouste, 2019; Calvo, Sarkisian, and Tamborini 2013; Gorry, Gorry, and Slavov 2018; Mazzonna and Peracchi, 2017). Specifically, we compared each group of retirees; that is, those who retired within five years and those who retired more than five years ago were compared with those who were still working. Appendix Table G.24 shows that retirement was associated with improved physical independence and self-rated health in both groups, whereas the associations between retirement and cognitive function were not statistically significant. Furthermore, we observed a significant association between retirement and decreased physical inactivity among those who retired more than five years ago but not among those who retired within five years.

Finally, given that marital status can affect the decision of retirement, we included interactions between IVs and marital status in the FEIV models. The first stage estimation in Appendix Table G.25 suggests that married people were more likely to retire when they reached SPA than those who were not married.[23] The second stage estimation in Appendix Table G.26 shows that our findings are robust even after the modification of the first stage.

2.5 Discussion and Conclusions

This is the first study to examine the impact of retirement on cognitive function, physical independence, and self-rated health using harmonized longitudinal data from 35 countries. Our FEIV models revealed that retirement was associated with improved cognitive function

---

[22] Imputing missing values using a hierarchical model is a common strategy in longitudinal studies. In this case, the imputation model used a linear time trend to account for changes over time and included several variables to predict missing values, such as sex, age, marital status, working status, retirement status, ERA and ORA status, three outcome variables, three health behavior variables, and country. Assuming missing at random, the imputation model estimates the missing values based on the available data and the variables that are predictive of the missing values.

[23] The first stage of the model for binge drinking did not show clear interactions between IVs and marital status, which could be attributable to small sample size and differences in available countries in the analysis.

and physical independence among women. In both sexes, retirement improved self-rated health, but women indicated a larger effect than men.

The sex difference in the effect of retirement on cognitive function has been reported in previous studies (Atalay et al., 2019; Ebeid & Oguzoglu, 2023); however, its detrimental effect among men was statistically insignificant in our study. The observed effect size of 0.10 SD among women was not negligible, given that Kraft (2020) proposed to consider the effect of 0.05–0.20 SD on cognition as a medium size.[24] Our stratified analysis of health behaviors showed that retirement was associated with decreased physical inactivity and smoking among women but not among men. Post-retirement health behaviors can be mechanisms through which retirement influences health (Eibich, 2015). Given that unhealthy lifestyles including physical inactivity and smoking are potential risk factors for cognitive and memory declines (Jia et al., 2023; Kesse-Guyot et al., 2014; Sabia et al., 2009), the sex differences in health behaviors may induce the heterogeneous effect of retirement on cognitive function.

Our findings on physical independence were contrary to non-IV studies that showed negative associations with retirement (Dave et al., 2008; Stenholm et al., 2014). As we demonstrated, the effect of retirement flipped to be positive after adopting the IV, which suggested that the previous studies could not fully address its endogeneity. This study was consistent with more recent studies (Szabó et al., 2019; van Zon et al., 2016). For example, adopting FEIV models, Nishimura, Oikawa, and Motegi (2018) showed the positive impact of retirement on ADL in England using data from the ELSA between 2002 and 2014, as well as in the United States, Germany, Switzerland, Japan, and South Korea. Similar to cognitive function, the estimates of physical independence were consistent with the associations of retirement with physical inactivity and smoking in favor of women.

Similar to many previous studies (Coe and Zamarro 2011; Eibich 2015; Gorry, Gorry, and Slavov 2018; Hessel 2016; Johnston and Lee 2009; Messe and Wolff 2019; Neuman, 2008; Nishimura, Oikawa, and Motegi 2018; Rose 2020; Zhu 2016), we found improved self-rated health among retirees compared with full-time workers. There are several possible pathways that link retirement to better self-rated health. Retirement provides opportunities for individuals to engage in health-promoting activities, including exercise, healthy eating, and adequate sleep (Barnett et al., 2012; Helldán et al., 2012; Kämpfen & Maurer, 2016; Myllyntausta et al., 2018). We observed that women exhibited a greater effect size on self-

[24] The author reviewed 1,942 effect sizes from 747 randomized control trials that evaluated the effect of educational interventions on cognitive skills. He found that the distribution of effect sizes had a median of 0.10 SD and even the 90th percentile was under 0.50 SD.

rated health than men, which may reflect their healthier lifestyles after retirement (i.e., decreased physical inactivity and smoking). Furthermore, retirees tend to spend more time in social activities that promote mental and physical well-being (Bogaard et al., 2014; Kobayashi et al., 2022). Relief from job strain can be another mechanism explaining the relationship between retirement and improved self-rated health (Lerner et al., 1994).

In line with the sex differences in the health outcomes, we observed the effects of retirement on reduced physical inactivity and smoking only among women. The decreased physical inactivity for women is consistent with a previous study showing that women are more socially active than men after retirement (Atalay et al., 2019). Female retirees participate in more social activities and have more channels of social support than men, which can lead to increased opportunities to engage in physical activity (Kikuchi et al., 2017; Ståhl et al., 2001). Additionally, we found that women who retired from high-control jobs experienced reduced smoking. Smoking is one of the stress-coping behaviors, and women tend to be more reactive to stress compared to men (Bale & Epperson, 2015; Schmaus et al., 2008). Jobs with a high level of discretion are often associated with greater responsibility and, consequently, increased stress. This may have led women to adopt smoking as a means to cope with stress. Furthermore, in high-control jobs where interactions with colleagues and clients are crucial, women working in predominantly male environments might have used smoking as a means of maintaining social connections. Retirement could have resulted in a loss of motivation to smoke for women who held positions of high discretion, thus leading to a decreased likelihood of smoking.

There were several limitations in this study. First, certain discrepancies across surveys were recognized despite the fact that field experts harmonized data (Shih, Lee, and Das 2012; Zamarro and Lee 2012; Lee, Phillips, and Wilkens 2021). Even though these discrepancies could lead to estimations being biased, some potential biases could be eliminated by including the country FEs. Second, measurement errors could occur because most of the measures were self-reported. Nonetheless, the performance of outcome measures has been validated (Bond et al., 2006; Brenowitz et al., 2014; Idler & Benyamini, 1997; Morris et al., 1989; Roehrig et al., 2007; Welsh et al., 1994). In addition, the straightforward inquiry into retirement status ensured face validity to measure individuals' recognition, which could induce behavioral adjustments (Eibich, 2015). Third, further studies are needed to determine the mechanism linking retirement to improved outcomes. Our study demonstrated sex differences in physical inactivity and smoking after retirement, which was consistent with the heterogeneous effect of retirement on health by sex. Nonetheless, mediation analysis is necessary to confirm the mechanism. In addition, there may be other factors such as sleep, diet, and social participation that were not provided in the harmonized data.

This study suggests that retirement benefits health, especially for women. While we observed sex heterogeneity, the effects of retirement on the outcomes appeared constant across different educational levels, pre-retirement job characteristics, and countries. Notably, increasing the SPA delays retirement timing and might dampen citizens' health. Nevertheless, promoting healthy behaviors such as engaging in physical activity and refraining from smoking can offset the potentially detrimental effects of delayed retirement and contribute to realizing healthy aging.

3. Heterogeneous Treatment Effect of Retirement on Cognitive Function[25]

3.1 Introduction

The cognitive health of older people is a global concern. In 2015, there were 47 million people with dementia, and this number is projected to increase 1.6-fold to 75 million by 2030 (World Health Organization, 2017). Given the rapid escalation in the demographic prevalence of people with dementia, the resulting economic burden borne by society is poised to be significant. This burden includes not only the direct costs attributable to medical interventions and long-term care but also the indirect costs experienced by informal caregivers, such as opportunity costs, alternative labor expenses, foregone earnings, and the psychosocial encumbrances experienced (Cimler et al., 2019; Hurd et al., 2013; Wimo et al., 2017; Wittenberg et al., 2019). The key approach to mitigating the increasing prevalence of dementia lies in the comprehensive elucidation of the causal relationship between dementia and the biological, epidemiological, and socioeconomic determinants that could potentially exert an impact on its pathogenesis.

Within the realm of economics, a subset of investigations has particularly centered on the potential ramifications of retirement behavior on dementia, although a consensus in findings has not yet been unequivocally established. An underpinning challenge in empirical studies resides in the endogenous nature of characterizing the retirement decision, a phenomenon encapsulated as the "healthy worker survivor effect" (Arrighi & Hertz-Picciotto, 1994). The phenomenon in which healthier people tend to sustain their employment results in a fundamental disparity in cognitive function between those in the workforce and retirees. This inconsistency stems from inherent dissimilarities in principles between the two groups. In the absence of a thorough resolution of endogeneity within an empirical model, retirement becomes erroneously linked to a decline in cognitive function. To address these endogeneity issues, empirical researchers often use the state pension age (SPA) to identify the causal effects of retirement on cognitive function, assuming that reaching the SPA exogenously increases the probability of retirement.[26] Retirement dramatically changes individuals' budget

[25] The results presented in this chapter have been presented as Sato, Noguchi, & Inoue (2023).
[26] Many developed countries are increasing SPA to accommodate the rapidly aging population (Organisation for Economic Co-operation and Development, 2021). For example, the United States has increased the SPA from 65 to 66 by 2009 and restarted increasing it to 67 by 2027 (Li, 2022). The United Kingdom continues to increase the SPA from 65 to 67 by 2028 and has a further plan to increase it to 68 (United Kingdom Government, 2014).

constraints and time allocation between labor and leisure, which affects the level of health attained (Grossman 1972). Thus, policymakers need to pay attention to the potential ramifications of delayed retirement due to increasing SPA based on individuals' characteristics.

This study aims to explore the heterogeneous treatment effect of retirement on cognitive function using the instrumental variable (IV) forests algorithm developed by Athey, Tibshirani, and Wager (2019). The basic idea of the IV forests estimation is that it is a combination of the generalized method of moments (GMM) and random forests. We used SPA as an IV for retirement, and the GMM produced IV estimates. Random forests can detect observations that have similar treatment effects. Hence, the IV forests calibrate the conditional average treatment effects based on the GMM localized by "similarity" weights derived from a random forests-based algorithm. This novel method has several advantages for the investigation of heterogeneous treatment effects. First, it is a data-driven, machine learning-based approach capable of unveiling concealed effect modifiers. Conventional research has often focused on a limited set of modifiers, assessing their effect heterogeneity through interaction terms and the stratification of analytical samples. IV forests diverge from conventional approaches by accommodating a wide array of potential covariates, Second, it substantially mitigates the risk of model misspecification primarily because of its nonparametric nature. This is a notable feature in studying the effects of retirement on health, given that a previous review highlighted the inconsistency in the existing literature, partly attributing it to issues with model specifications (Nishimura et al., 2018). Third, the algorithm is superior to classical random-forest-based algorithms by providing asymptotic normal estimates using a sample splitting technique referred to as "honesty." This is an essential property for testing hypotheses and calculating confidence intervals.

Indeed, through its pioneering and initial application of the IV forest methodology, distinguished by its distinctive attributes described in the previous paragraph, this study on the impact of retirement on cognitive function makes noteworthy contributions to the extant research landscape, as follows. First, through a data-driven, machine-learning-based approach, this study has unveiled hitherto unacknowledged factors that modify the effect of retirement, including variables such as income, assets, and pre-retirement health conditions and behaviors. Retirees find themselves endowed with more leisure time but are constrained by financial resources for health-related investments. However, individuals with higher socioeconomic status during their pre-retirement phase possess the financial means to allocate resources toward enhancing their cognitive function. This can be understood within the framework of the Grossman model (Grossman 1972), in which individuals with greater financial capacity experience fewer constraints on their lifetime budgets than those with lower socioeconomic status. Similarly, our findings indicate that individuals with better health prior to retirement tend to exhibit superior cognitive function post-retirement, aligning with the theory that states that individuals with less time spent sick have the luxury of dedicating more time to health-related investments than their less-healthy counterparts. Consequently, we found that individuals with robust health, along with higher educational attainment, financial assets, and income, tend to accrue more substantial cognitive benefits from retirement.

Second, our study differs from prior studies as it refrains from imposing parametric assumptions and exhibits reduced susceptibility to model misspecification. This is a hallmark of the IV forests approach. As an illustration, Nishimura, Oikawa, and Motegi (2018) have previously elucidated that variances in outcomes within earlier research endeavors might stem from variations in model specifications, as observed in their replication of prior studies. The existing literature has shown inconsistent results even though researchers have used the same datasets and employed SPA as an IV for retirement. Using data from the Health and Retirement Study (HRS), Bonsang, Adam, and Perelman (2012) demonstrated the detrimental effect of retirement on cognitive function, whereas Coe et al. (2012) found no evidence of the effect. Among studies using data from the Survey of Health, Ageing and Retirement in Europe (SHARE), Celidoni, Dal Bianco, and Weber (2017) showed a detrimental effect, Coe and Zamarro (2011) indicated a non-significant association, and Bianchini and Borella (2016) found that retirement improved cognitive function. In other countries, studies using data from the English Longitudinal Study on Ageing (ELSA) and the Japanese Study of Aging and Retirement (JSTAR) did not find clear associations between retirement and cognitive function (Nishimura et al., 2018; Rose, 2020), whereas analysis using data from the Korean Longitudinal Study of Aging indicated a beneficial effect (Nishimura et al., 2018).

Finally, it is worth noting that our findings have significant and valuable policy implications, transcending the application of a new analytical framework for assessing the impact of retirement on cognitive function. Building upon the findings that postponing retirement could accelerate the decline in cognitive function in some individuals, we estimated the fiscal costs of dementia care resulting from an increase in SPA. Our projections indicate that the United Kingdom is poised to incur greater financial burdens than the United States due to the absence of early retirement options, thereby affecting a substantial portion of the workforce. The introduction of early retirement into the present system could, to some extent, alleviate escalating costs. Thus, we recommend that policymakers consider incorporating provisions for early retirement into the pension system to enable individuals to make retirement decisions according to their unique circumstances. Additionally, we underscore the favorable impact of physical activity on the post-retirement period. The promotion of physical activity initiatives can potentially alleviate the adverse effects of delayed retirement on cognitive health.

The remainder of this paper is organized as follows. Section 3.2 describes the data used in this study, Section 3.3 presents the empirical model, Section 3.4 reports the results, and Section 3.5 discusses the results and concludes the paper.

3.2 Data

3.2.1 Harmonized Panel Data

This study uses harmonized panel datasets from the HRS, ELSA, and SHARE provided by the Gateway to Global Aging Data project (Lee, Phillips, and Wilkens 2021).[27] Our data comprised three waves: we obtained covariates (except for age) from the HRS and ELSA in 2014 and SHARE in 2015; age and labor force status were ascertained via the HRS and ELSA in 2016 and SHARE in 2017; and the outcome of cognitive function was assessed in the HRS and ELSA in 2018 and SHARE in 2019.[28]

Appendix Figure H.1 presents a sample flowchart. Of the 94,824 individuals who participated in the first wave, 49,555 were followed-up with in all three waves. We included 43,052 individuals aged 50–80 years in the second wave but excluded 29,519 individuals who did not work in the first wave and 722 individuals who neither worked nor retired in the second wave (e.g., unemployed, disabled, or homemaker). Finally, 12,811 participants were included in the development of IV forests.

[27] The harmonized datasets are available from https://g2aging.org/ (Accessed: January 21, 2023). This project provides "a free public resource designed to facilitate cross-national and longitudinal studies on aging." Although the harmonized datasets of the Irish Longitudinal Study on Ageing, the Longitudinal Aging Study in India, and the Malaysia Ageing and Retirement Study were also available, they were harmonized only for one wave and thus excluded. Data from the Mexican Health and Aging Study and the China Health and Retirement Longitudinal Study were also excluded because they conducted interviews triennially and their harmonized variables were limited. We neither used data from the Costa Rican Longevity and Healthy Aging Study, the Japanese Study of Aging and Retirement, and the Korean Longitudinal Study of Aging because we previously found that IVs in these countries were weak (Sato & Noguchi, 2023).

[28] A total of 17 countries participated in all the three waves of the SHARE, namely, Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, France, Germany, Greece, Israel, Italy, Luxembourg, Poland, Slovenia, Spain, Sweden, and Switzerland.

3.2.2 Cognitive Function

We examined episodic memory as a measure of cognitive function. It involves the ability to recall past experiences, which declines with age (Tulving 2002). It was assessed in accordance with the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Battery (Morris et al., 1989). Participants listened to 10 common words and were immediately asked by an interviewer to recall as many words as possible. They were then asked to recall the words again after approximately five minutes. Hence, the total number of words that the participants could recall ranged between 0 and 20 and represented their cognitive function, as in previous studies (Bianchini & Borella, 2016; Bonsang et al., 2012; Celidoni et al., 2017; Coe et al., 2012; Coe & Zamarro, 2011; Nishimura et al., 2018; Rose, 2020).

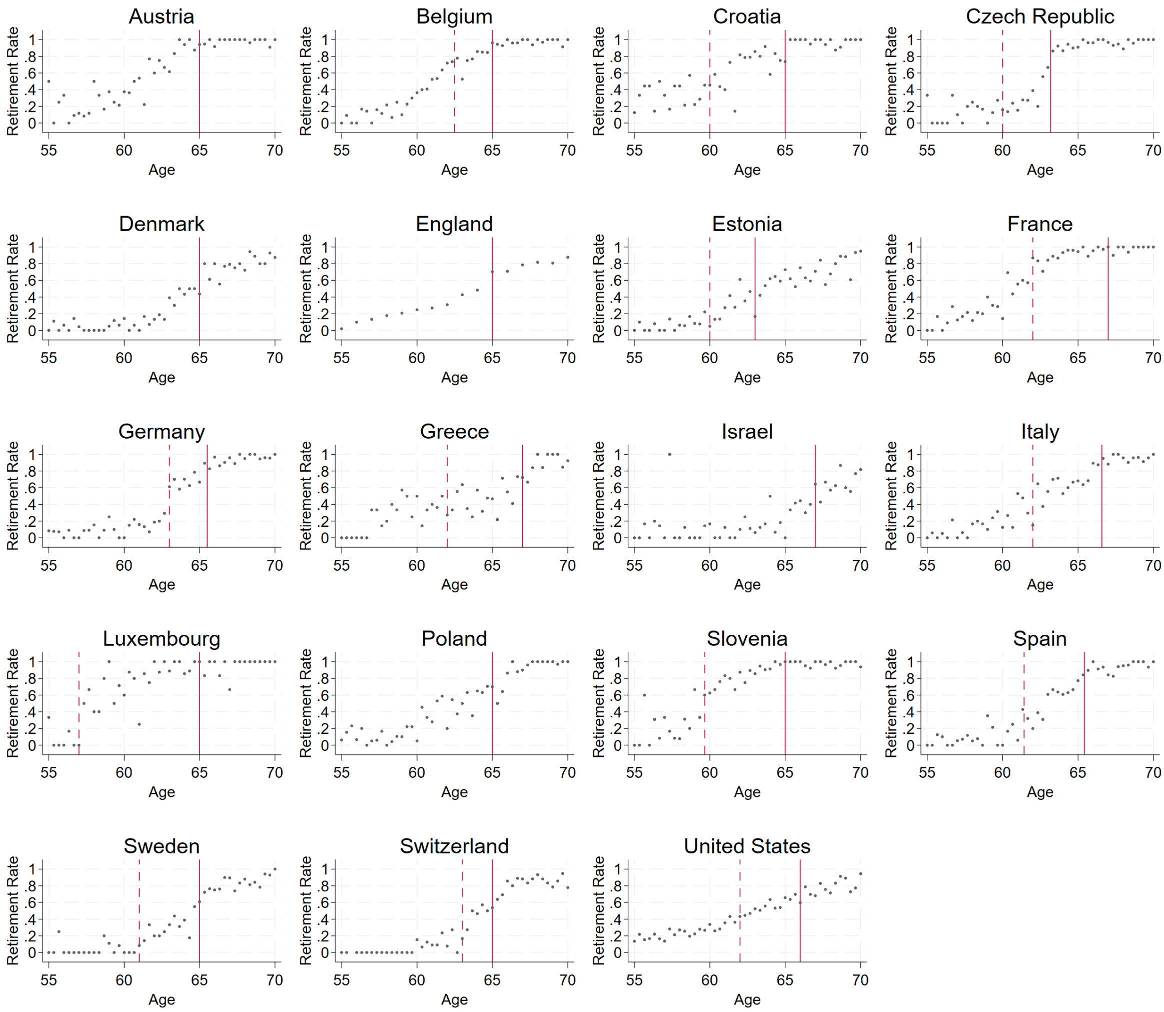
3.2.3 Retirement and State Pension Age

Labor force status was self-reported in the surveys, as described in Appendix C and by Zamarro and Lee (2012). We restricted the sample to those who worked during the first wave. In the second wave, we defined retirees as those who self-identified as "retired," regardless of their working status (i.e., including those who were "partly retired"), following previous literature (Atalay et al., 2019; Bianchini & Borella, 2016). Other studies have defined retirement as not working (Bingley & Martinello, 2013; Bonsang et al., 2012; Coe & Zamarro, 2011), and we checked the robustness of our findings using a narrower definition of retirement.

To eliminate bias stemming from endogenous selection for retirement, SPA was used as an IV for retirement. We employed the joint instruments of early retirement age (ERA) and official retirement age (ORA) to predict retirement following the method of a previous study (Coe & Zamarro, 2011). A binary ERA variable discerned whether participants reached the earliest eligibility age for receiving either reduced or full pension, subject to specific conditions. Likewise, a binary ORA variable denoted whether participants reached the age of entitlement to the minimum guaranteed pension or full pension without any requirements. The ERA variable was set to zero for all participants in countries where early retirement schemes were not implemented. Appendix D shows the SPA of each country collected from "Social Security Programs Throughout the World" (United States Social Security Administration, 2020), "Pensions at a Glance" (Organisation for Economic Co-operation and Development, 2021), and websites of the national authorities. Figures 3.1 (men) and 3.2

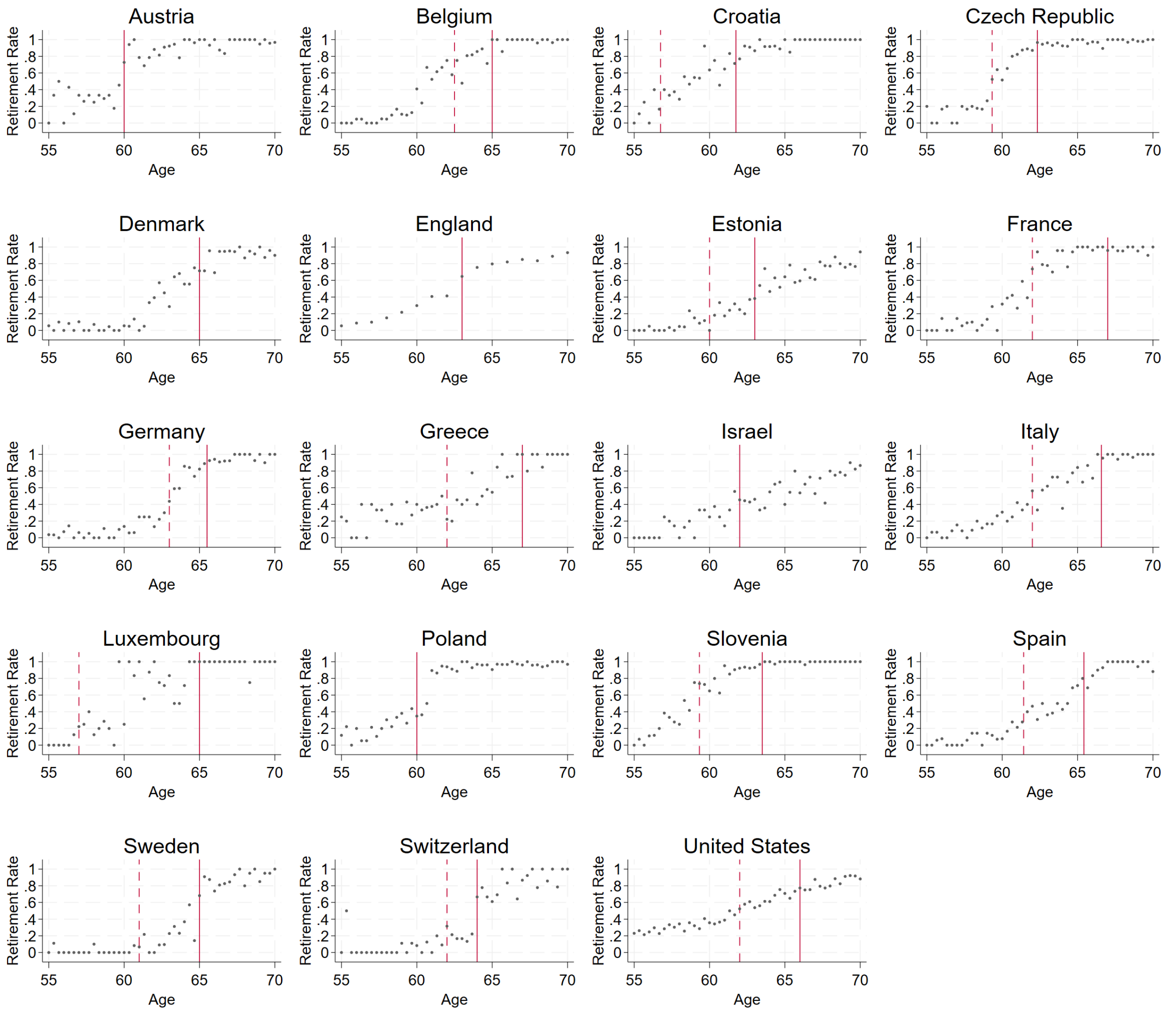
(women) describe the changes in retirement rates according to age using data from the second wave. We observed some jumps in retirement rates around the SPA.

Figure 3.1 Retirement Rate of Men



Each dot represents the average retirement rate for each 4-month interval (monthly age data was unavailable in England). The retirement rate is calculated by dividing the number of retirees by the sum of retirees and workers. The dashed red line denotes the early retirement age, while the solid red line represents the official retirement age in the year of the second wave survey.

Figure 3.2 Retirement Rate of Women



Each dot represents the average retirement rate for each 4-month interval (monthly age data was unavailable in England). The retirement rate is calculated by dividing the number of retirees by the sum of retirees and workers. The dashed red line denotes the early retirement age, while the solid red line represents the official retirement age in the year of the second wave survey.

3.2.4 Covariates

We considered 60 harmonized covariates obtained from the first wave to develop IV forests. Table 3.1 presents definitions of these covariates.

Table 3.1 Definition of Covariates

Covariate	Type	Definition
Age	Continuous	This variable is the participant's age in months at the time of the second wave interview.
Men	Binary	This variable is coded as 1 for men and 0 for women.
Foreign-born	Binary	This variable is coded as 1 if the interview did not take place in the country of birth and 0 otherwise.
Education	Ordered	This variable is coded as 1 for less than upper secondary education, 2 for upper secondary and vocational training, and 3 for tertiary education according to the 1997 International Standard Classification of Education.
Married	Binary	This variable is coded as 1 for married or partnered and 0 for otherwise.
Living alone	Binary	This variable is coded as 1 for those whose household size is 1 and 0 for otherwise.
Number of children	Binary	Based on a variable indicating the number of participant's living children (including natural, foster, adopted, or stepchildren), we created two binary variables; "no children" indicates 1 if the number of children is zero and 0 for otherwise; "≥3 children" indicates 1 if the number of children is three or more and 0 for otherwise.
Asset	Continuous	This variable is the net value of assets at the couple-level unit calculated as the value of all wealth components (including housing, financial, and non-financial assets) minus that of all debts. To make the variables in different surveys comparable, we standardized them to z-scores for each survey. See Angrisani & Lee (2012b) for details about the harmonization of wealth measures.
Income	Continuous	This variable is the total income at the couple-level including earnings, capital income, pensions, and public transfers. To make the variables in different surveys comparable, we standardized them to z-scores for each survey. See Angrisani & Lee (2012a) for details about the harmonization of income measures.
Occupation	Binary	We created five binary variables indicating the participant's occupation: professional, clerk, service and sales, and manual labor. We categorized occupations based on the 2010 Census occupations in the HRS, the Standard Occupational Classification (2000) in the ELSA, and the 1988 International Standard Classification of Occupations in the SHARE. See Appendix Table H.1 for details about the occupational codes.
Physical demand	Ordered	This variable is a 4-point Likert scale indicating the degree to which the participant agrees that their job is physically demanding: 1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree.
Part-time job	Binary	This variable is coded as 1 if the participant works less than 35 hours per week and 0 otherwise.
Self-employed	Binary	This variable is coded as 1 if the participant reports to be self-employed and 0 for otherwise.
Baseline cognition	Continuous	This variable indicates baseline cognitive function measured in the same way as the outcome.
Self-rated health	Ordered	This variable is a 5-point Likert scale indicating self-rated health: 1 = poor; 2 = fair; 3 = good; 4 = very good; 5 = excellent.
Depression	Continuous	Higher scores of this variable indicate more severe depression. The HRS and the ELSA use a short version of the Center for Epidemiologic Studies Depression (CES-D) to measure depression, while the SHARE uses the EURO-D scale. To make the variables using different measures comparable, we standardized them to z-scores for each survey.

Life satisfaction	Continuous	Higher scores of this variable indicate higher levels of the participant's life satisfaction. The HRS uses a 5-point Likert scale, the ELSA uses a 7-point Likert scale, and the SHARE uses a 10-point Likert scale to measure life satisfaction. The harmonized datasets provide a variable standardized to z-scores for each survey to make them comparable.
Diagnosed diseases	Binary	We have nine variables of chronic medical conditions: namely, hypertension, diabetes, cancer, lung disease, heart disease, stroke, arthritis, psychiatric problems, and hyperlipemia. These variables indicate 1 if a doctor has ever told the participant that he or she has the conditions and 0 for otherwise. See Hu & Lee (2012) for details about the harmonization of chronic medical conditions.
Health limitation	Binary	This variable indicates 1 if the participant reports that an impairment or health problem limits the kind or amount of paid work and 0 for otherwise.
Difficulty in ADL	Binary	This variable indicates 1 if the participant has difficulties with any of the five ADL including bathing or showering, dressing, eating, getting in and out of bed, and walking across a room, and 0 for otherwise.
Difficulty in IADL	Binary	This variable indicates 1 if the participant has difficulties with any of the five IADL including using the telephone, managing money, taking medications, shopping for groceries, and preparing a hot meal, and 0 for otherwise.
Eyesight and hearing	Ordered	We have three 5-point Likert scales for self-reported distance eyesight, near eyesight, and hearing: 1 = poor; 2 = fair; 3 = good; 4 = very good; 5 = excellent.
Pain problems	Binary	This variable indicates 1 if the participant reports to be troubled with pain and 0 for otherwise.
Obesity	Binary	This variable indicates 1 if the participant's body mass index is 30 kg/m^2 or higher and 0 otherwise (World Health Organization, 2021).
Physical activity	Binary	This variable indicates 1 if the participant engages in vigorous or moderate physical activity once or more per week and 0 for otherwise.
Heavy drinking	Binary	This variable indicates 1 if the participant reports having 15 or more drinks per week for men and 8 or more drinks for women, and 0 for otherwise (National Institute on Alcohol Abuse and Alcoholism, 2023).
Smoking	Binary	This variable indicates 1 if the participant reports smoking now and 0 for otherwise.
Countries	Binary	We have 19 binary variables indicating the place of the interview: namely, Austria, Belgium, Croatia, Czech Republic, Denmark, England, Estonia, France, Germany, Greece, Israel, Italy, Luxembourg, Poland, Slovenia, Spain, Sweden, Switzerland, and United States.

3.3 Empirical Model

3.3.1 IV forests

To assess the heterogeneous treatment effect of retirement, we used an IV forests algorithm developed by Athey, Tibshirani, and Wager (2019). Suppose that n samples indexed by $i = 1, \dots, n$ are independent and identically distributed. Observations $O_i = \{Y_i, W_i, Z_i\}$ include an outcome $Y_i \in \mathbb{R}$ (cognitive function), a treatment assignment $W_i \in \{0,1\}$ (retirement), and an IV $Z_i \in \{0,1\}$ (SPA), along with a set of auxiliary covariates $X_i \in \mathcal{X}$. The conditional effects of interest $\theta(x)$ are solutions to the local moment conditions

$$\mathbb{E}\left[\psi_{\theta(x),\upsilon(x)}(O_i) \,\middle|\, X_i = x\right] = 0 \quad \forall\, x \in \mathcal{X},$$

where $\psi(\cdot)$ is a scoring function and $\nu(x)$ is an optional nuisance parameter. The IV forests estimates $\hat{\theta}(x), \hat{\nu}(x)$ are obtained by solving

$$(\hat{\theta}(x), \hat{\nu}(x)) \in \arg\min_{\theta,\nu} \left\{ \left\| \sum_{i=1}^{n} \alpha_i(x) \psi_{\theta,\nu}(O_i) \right\|_2 \right\}.$$

If the expression has a unique root, $(\hat{\theta}(x), \hat{\nu}(x))$ solves $\sum_{i=1}^{n} \alpha_i(x) \psi_{\hat{\theta}(x),\hat{\nu}(x)}(O_i) = 0$. The IV forests incorporate similarity weights $\alpha_i(x)$ to solve the heterogeneous estimating equation. The weights $\alpha_i(x)$ are obtained using random forests with a set of B trees indexed by $b = 1, \ldots, B$

$$\alpha_{bi}(x) = \frac{1(\{X_i \in L_b(x)\})}{|L_b(x)|}, \qquad \alpha_i(x) = \frac{1}{B} \sum_{b=1}^{B} \alpha_{bi}(x).$$

$L_b(x)$ denotes the set of training samples falling in the same leaf as a target sample x in tree b, and $\alpha_i(x)$ represents how often the ith training sample falls into the same leaf as x. The forests-based algorithm splits training samples to maximize the squared difference in treatment effect estimates across leaves (i.e., heterogeneity) so that $\alpha_i(x)$ leads to a good fit of $\theta(x)$. The estimates $\hat{\theta}_x$ have the property of asymptotic normality using a subsampling technique called "honesty" (Athey et al., 2019; Athey & Imbens, 2016; Wager & Athey, 2018). The basic idea of "honesty" is to divide the sample into three subsets; the "splitting" subset is used to partition samples and develop a tree; the "estimation" subset is used to estimate a treatment effect for each leaf of the fitted tree; and the "test" subset is used to validate the estimates.

To apply the forests-based algorithm to an IV regression, Athey, Tibshirani, and Wager (2019) assumed a structural model

$$Y_i = \mu(X_i) + \tau(X_i)W_i + \varepsilon_i$$

where $\mu(X_i)$ denotes a nuisance intercept parameter, $\tau(X_i)$ is interpreted to be the causal effect of W_i on Y_i, and ε_i is an error term that can be correlated with W_i. To recover the consistency of $\tau(X_i)$ in the case of a correlation between W_i and ε_i, an IV Z_i is used. If Z_i is independent of ε_i conditionally on X_i, and the covariance of Z_i and W_i conditionally on X_i is nonzero, $\tau(X_i)$ is identified as

$$\tau(X_i) = \frac{\mathrm{Cov}[Y_i, Z_i \mid X_i = x]}{\mathrm{Cov}[W_i, Z_i \mid X_i = x]}.$$

In this setting, a scoring function $\psi(\cdot)$ can be defined as

$$\psi_{\tau(x),\mu(x)}(Y_i, W_i, Z_i) = \begin{bmatrix} Z_i(Y_i - W_i\tau(x) - \mu(x)) \\ Y_i - W_i\tau(x) - \mu(x) \end{bmatrix}.$$

Then, $\tau(x)$ is estimated via moment functions $\mathbb{E}\left[Z_i\left(Y_i - W_i\tau(x) - \mu(x)\right)\middle|X_i = x\right] = 0$ and $\mathbb{E}[Y_i - W_i\tau(x) - \mu(x)|X_i = x] = 0$. Biewen and Kugler (2021) extended IV forests to a multiple IVs setting (i.e., Z_i is a $M \times 1$ vector) by defining $\psi(\cdot)$ as

$$\psi_{\tau(x),\mu(x),\gamma_1(x),\gamma_0(x)}(Y_i, W_i, Z_i) = \begin{bmatrix} \dot{W}_i(Y_i - \dot{W}_i\tau(x) - \mu(x)) \\ Y_i - \dot{W}_i\tau(x) - \mu(x) \\ Z_i(W_i - Z_i'\gamma_1(x) - \gamma_0(x)) \\ W_i - Z_i'\gamma_1(x) - \gamma_0(x) \end{bmatrix},$$

where $\dot{W}_i = \gamma_0(x) + Z_i'\gamma_1(x)$. The estimates of a conditional local average treatment effect $\hat{\tau}(x)$ can be obtained by solving $M + 3$ moment conditions, along with weights $\alpha_i(x)$.

3.3.2 Statistical Analysis

Our IV forests developed 2000 trees and tuned their parameters[29] through cross-validation. To summarize the effect of retirement on cognitive function obtained using IV forests, we estimated the local average treatment effect on the overlap population (LATO). We define the conditional mean of the outcome as $y(x) = \mathbb{E}[Y_i|X_i = x]$, the propensity score of the treatment as $w(x) = \mathbb{E}[W_i|X_i = x]$, and the propensity score of the instrument as $z(x) = \mathbb{E}[Z_i|X_i = x]$. The IV forests yield out-of-bag estimates (i.e., the prediction of the ith observation is obtained via trees fitted without using the ith observation) of these marginal expectations, $\hat{y}^{(-i)}, \hat{w}^{(-i)}, \hat{z}^{(-i)}$, which recover $\sqrt{n}$ consistency of estimates using a machine learning-based method (Chernozhukov et al., 2018). Then, we limited samples to the overlap population $\mathcal{P}$ by trimming the estimated propensity score of treatment $\hat{w}^{(-i)}$ to a value between 0.1 and 0.9 (Crump et al., 2009). To obtain the LATO relying on the non-parametric estimation, we computed conditionally centered outcomes $\tilde{Y}_i = Y_i - \hat{y}^{(-i)}(X_i)$, $\widetilde{W}_i = W_i -$

[29] Namely, the fraction of the data used to build each tree, the number of variables tried for each split, a target for the minimum number of observations in each leaf, the fraction of data that will be used for determining splits, whether to prune the estimation sample tree such that no leaves are empty, a parameter that controls the maximum imbalance of a split, and a parameter that controls how harshly imbalanced splits are penalized.

$\hat{w}^{(-i)}(X_i)$, and $\tilde{Z}_i = Z_i - \hat{z}^{(-i)}(X_i)$, and then ran residual-on-residual two-stage least squares (2SLS) regression using centered outcomes $\{\tilde{Y}_i, \tilde{W}_i, \tilde{Z}_i\}_{i=1}^{n} \in \mathcal{P}$.[30] For the purpose of comparison, we also performed parametric ordinary least squares (OLS) and 2SLS regressions adjusting for 10 covariates selected based on variable importance in trained IV forests (i.e., covariates most frequently used to split samples) among the overlap population. In addition, we ran causal forests without an IV.

To see how well our IV forests captures effect heterogeneity, we drew a calibration plot according to the ranking of the estimated conditional local average treatment effect on the overlap population (CLATO) $\hat{\tau}(x)$. To obtain a valid inference of $\hat{\tau}(X_i)$ for the ith observation based on the "honesty" property, we should fit trees without using the ith observation. Hence, we divided the sample into 10 folds. Then, IV forests were fitted using nine folds, and the remaining fold was used to predict $\hat{\tau}(x)$. According to the quintile of the $\hat{\tau}(x)$ ranking within the held-out fold, we categorized observations into five groups from Q1 (the lowest CLATO; the subgroup of individuals who received the least benefits from retirement) to Q5 (the highest CLATO; the subgroup of individuals who received the most benefits from retirement). This procedure was repeated for each iteration.[31] Finally, we estimated LATOs for each quintile subgroup. Furthermore, to assess the heterogeneity of each covariate, we compared the mean values of covariates across groups. For continuous variables, we also depicted a partial dependence plot with the continuous variable on the x-axis and the out-of-bag prediction of CLATO $\hat{\tau}^{(-i)}(X_i)$ on the y-axis.

Building on the LATO estimates for the quintile groups, we estimated the monetary costs of dementia care induced by increasing ORA from the age of 65 to 66 years in the United States and the United Kingdom.[32] We relied on a previous study that indicated that a one-word increase in the word recall test predicts 0.85 times lower odds of dementia in five years (Tierney et al., 2010). Additionally, we predicted increases in the number of workers aged 66 by estimating the reduced probabilities of retirement attributable to increases in ORA using HRS and ELSA samples. Hurd et al. (2013) and Wittenberg et al. (2019) provided monetary cost estimates for dementia care per patient and their projections for 2030 in the

[30] We did not simply average $\hat{\tau}(x)$ to obtain LATO because individual predictions of $\hat{\tau}(x)$ have large errors and cannot estimate LATO efficiently. In contrast, Robinson (1988) showed that this orthogonal transformation yields $\sqrt{n}$ consistent estimates, even if nuisance estimates such as $y(x), w(x), z(x)$ converge at a slower rate.
[31] Namely, we performed a 10-fold cross-fitting procedure.
[32] The United States and the United Kingdom increased their ORA to age 66 for those born after 1943 and those born after 1954, respectively. Both countries have plans to further increase their ORA to age 67.

United States and the United Kingdom, respectively. By multiplying these components, we estimated the changes in the total cost of dementia care in 2030.

In all analyses, missing values were imputed using a random forests-based algorithm (Mayer, 2021)[33], assuming that the data were missing at random. Appendix Table H.2 reports the imputed values for each variable.

3.4 Results

3.4.1 Descriptive Statistics

After trimming the samples based on the propensity score for retirement, our analytical sample comprised 7,432 individuals, including 5,267 (70.9%) workers and 2,165 (29.1%) retirees in the second wave. Table 3.2 presents descriptive statistics. Workers had higher cognitive function than retirees in the third wave. In the first wave, we found unbalanced characteristics even though all of them had been working. Those who continued working were younger and had a higher education and were more likely to be foreign-born, professional workers, and self-employed at baseline than those who retired in the second wave. However, workers were less likely to engage in manual labor and part-time jobs than retirees. Furthermore, workers had higher baseline cognitive function, self-rated health, and hearing ability and were less likely to have health problems, including hypertension, diabetes, lung disease, stroke, arthritis, health limitations in working, difficulties in activities of daily living, pain-related problems, and smoking habits, than retirees. We also found differences between workers and retirees in the composition of countries. The outcomes appeared to be normally distributed, as shown in Appendix Figure H.2.

Table 3.2 Descriptive Statistics of the Overlap Population

Variables, n (%)	Worker n = 5267 (70.9%)	Retiree n = 2165 (29.1%)	P-value
Outcome			
Cognitive function, mean (SD)	11.1 (3.21)	10.8 (3.19)	<0.001
Characteristics			
Age, year, mean (SD)	63.9 (3.86)	65.8 (4.27)	<0.001
Men	2667 (50.6)	1063 (49.1)	0.23
Foreign-born	707 (13.4)	230 (10.6)	<0.001
Education, mean (SD)	2.2 (0.68)	2.1 (0.69)	<0.001
Married	4059 (77.1)	1663 (76.8)	0.81
Living alone	934 (17.7)	406 (18.8)	0.30

[33] We set the number of candidate non-missing values to sample from in the predictive mean matching steps to 3 and the number of trees to 100.

No children	544 (10.3)	200 (9.2)	0.15
≥3 children	1885 (35.8)	760 (35.1)	0.58
Asset, z-score, mean (SD)	0.1 (1.21)	0.0 (0.79)	0.48
Income, z-score, mean (SD)	0.1 (1.14)	0.0 (0.91)	0.67
Professional	2272 (43.1)	878 (40.6)	0.04
Clerk	793 (15.1)	329 (15.2)	0.88
Service & sales	1067 (20.3)	446 (20.6)	0.74
Manual labor	1140 (21.6)	515 (23.8)	0.04
Physical demand, mean (SD)	2.3 (1.07)	2.3 (1.05)	0.91
Part-time job	1298 (24.6)	775 (35.8)	<0.001
Self-employed	1151 (21.9)	368 (17.0)	<0.001
Health & Behaviors			
Baseline cognition, mean (SD)	11.3 (3.19)	11.0 (3.26)	<0.001
Self-rated health, mean (SD)	3.4 (0.98)	3.3 (0.95)	<0.001
Depression, z-score, mean (SD)	0.0 (1.01)	0.0 (0.97)	0.58
Life satisfaction, z-score, mean (SD)	0.0 (1.00)	0.1 (0.98)	0.06
Hypertension	2064 (39.2)	982 (45.4)	<0.001
Diabetes	626 (11.9)	311 (14.4)	0.003
Cancer	372 (7.1)	177 (8.2)	0.10
Lung disease	231 (4.4)	122 (5.6)	0.02
Heart disease	571 (10.8)	265 (12.2)	0.08
Stroke	107 (2.0)	69 (3.2)	0.003
Arthritis	1633 (31.0)	800 (37.0)	<0.001
Psychiatric problems	542 (10.3)	244 (11.3)	0.21
Hyperlipemia	1534 (29.1)	634 (29.3)	0.89
Health limitations in working	526 (10.0)	291 (13.4)	<0.001
Difficulty in ADL	205 (3.9)	137 (6.3)	<0.001
Difficulty in IADL	129 (2.4)	45 (2.1)	0.34
Distance eyesight, mean (SD)	3.8 (0.94)	3.8 (0.92)	0.96
Near eyesight, mean (SD)	3.6 (0.98)	3.6 (0.97)	0.21
Hearing, mean (SD)	3.6 (1.00)	3.5 (0.99)	0.01
Pain problems	1702 (32.3)	783 (36.2)	0.001
Obesity	1515 (28.8)	612 (28.3)	0.67
Physical activity	4629 (87.9)	1888 (87.2)	0.42
Heavy drinking	524 (9.9)	237 (10.9)	0.20
Smoking	795 (15.1)	369 (17.0)	0.04
Countries			
Austria	50 (0.9)	37 (1.7)	0.006
Belgium	99 (1.9)	58 (2.7)	0.03
Croatia	37 (0.7)	14 (0.6)	0.79
Czech Republic	147 (2.8)	126 (5.8)	<0.001
Denmark	291 (5.5)	92 (4.2)	0.02
Estonia	379 (7.2)	93 (4.3)	<0.001
France	123 (2.3)	86 (4.0)	<0.001
Germany	262 (5.0)	128 (5.9)	0.10
Greece	182 (3.5)	49 (2.3)	0.007
Israel	100 (1.9)	30 (1.4)	0.13
Italy	121 (2.3)	37 (1.7)	0.11

Luxembourg	44 (0.8)	27 (1.2)	0.10
Poland	31 (0.6)	13 (0.6)	0.95
Slovenia	94 (1.8)	60 (2.8)	0.007
Spain	108 (2.1)	61 (2.8)	0.04
Sweden	263 (5.0)	152 (7.0)	<0.001
Switzerland	276 (5.2)	115 (5.3)	0.90
England	974 (18.5)	403 (18.6)	0.90
United States	1686 (32.0)	584 (27.0)	<0.001

Imputed data are used.

3.4.2 Average Treatment Effects

Table 3.3 compares the LATO estimated using IV forests with the estimates of the conventional OLS, 2SLS, and non-IV forests. OLS and 2SLS models were adjusted for 10 covariates selected based on variable importance[34] in trained IV forests: assets, age, income, baseline cognition, depression, life satisfaction, self-rated health, hearing, degree of physical demands of the job, and distance eyesight (see Appendix Figure H.3 for the variable importance of each covariate). The OLS test showed a non-significant negative association between retirement and cognitive functioning. The first-stage estimates of the 2SLS indicated that reaching the ERA and ORA significantly increased the probability of retirement. The F-statistic showed a strong correlation between IVs and retirement, and the over-identification test did not reject the null hypothesis, suggesting that our IVs were valid. In the second-stage estimates, retirement was significantly associated with increased cognitive function. Similar to the OLS results, non-IV forests showed a non-significant negative association. However, the IV forests indicated that retirees could recall 1.348 more words than workers, and the point estimate was statistically significant.

Table 3.3 Average Treatment Effects of Retirement on Cognitive Function

	(1) OLS[a]	(2) 2SLS[a]	(3) Non-IV Forests	(4) IV Forests
Retirement	-0.013	0.962***	-0.031	1.348**
	(0.070)	(0.344)	(0.071)	(0.528)
ERA (1st stage)		0.091***		
		(0.011)		
ORA (1st stage)		0.249***		
		(0.015)		
Observations	7432	7432	7432	7432
R squared	0.313	0.295	0.000	-0.051

[34] Variable importance is a simple weighted sum of how many times each covariate was split on at each depth in the forest. It should be noted that it indicates frequency with which the covariate was used for prediction and does not evaluate the strength of causal effect modification.

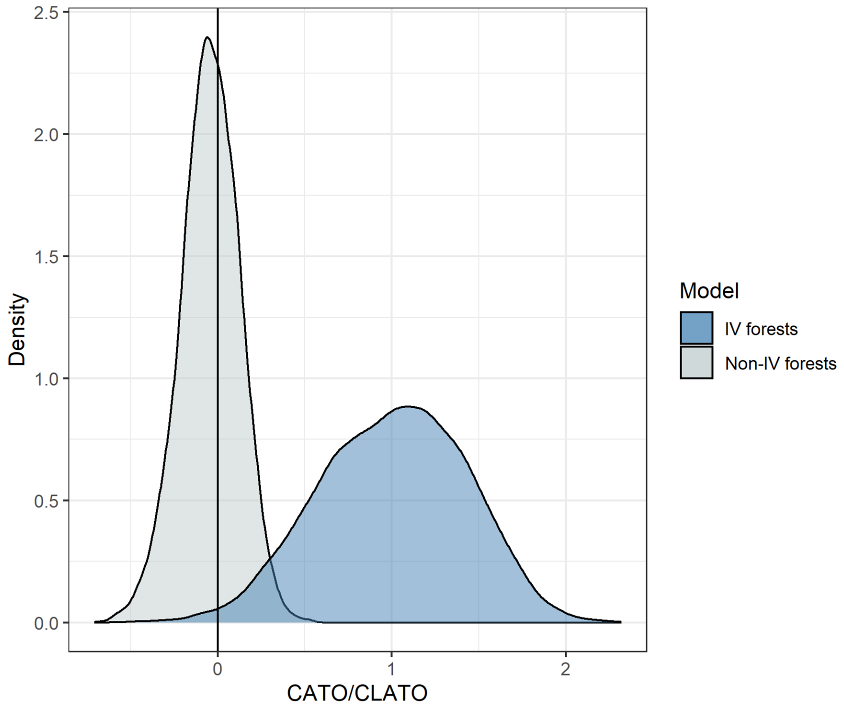
F statistic	163.037***
Sargan statistic	1.177

Standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
[a] The model is adjusted for the following covariates selected based on variable importance in trained IV forests: assets, age, income, baseline cognition, depression, life satisfaction, self-rated health, hearing, degree of physical demands of the job, and distance eyesight.

3.4.3 Heterogeneous Treatment Effect of Retirement

Figure 3.3 shows the distribution of the conditional average treatment effects comparing the non-IV and IV forests. While the estimates for non-IV forests were concentrated around zero, those for IV forests appeared to be heterogeneously distributed.

Figure 3.3 Distribution of Conditional Average Treatment Effects



The estimand of non-IV forests is the conditional average treatment effect on the overlap population (CATO), while that of IV forests is the conditional local average treatment effect on the overlap population (CLATO).

Figure 3.4 shows the calibration plot for CLATO. As the CLATO ranking increased, the estimated LATOs in these categories increased monotonically, suggesting that our IV forests correctly captured the heterogeneity in the effect of retirement on cognitive function. The point estimate of retirement indicated a harmful effect on cognitive function in the lowest CLATO group (Q1), whereas it showed protective effects in the other groups, although the 95% confidence interval included a value of 0.

Figure 3.4 Calibration Plot for CLATO

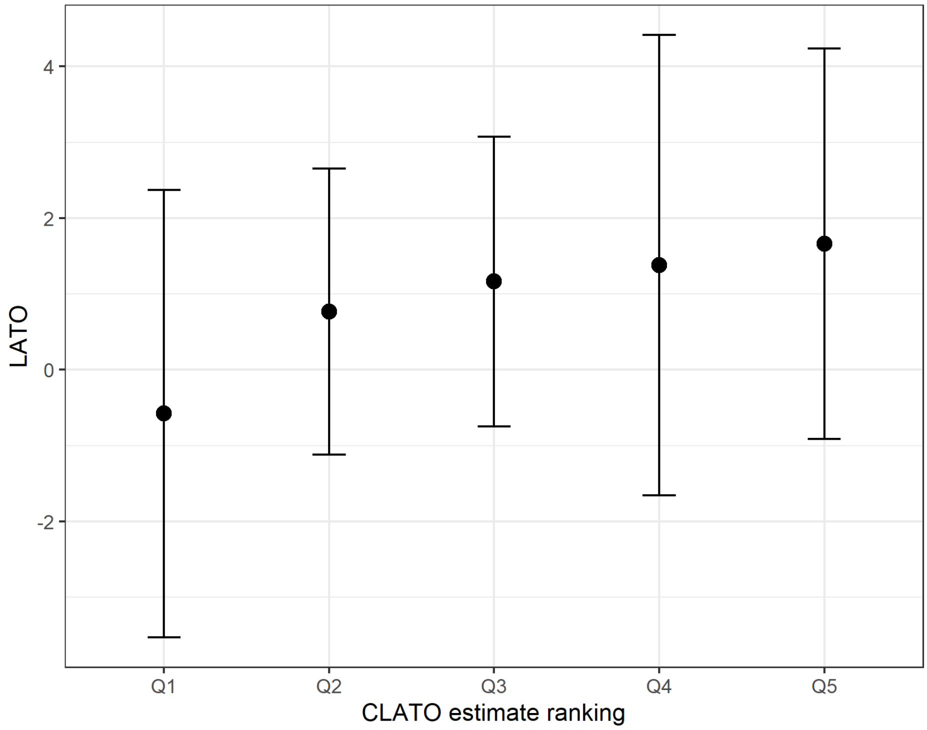


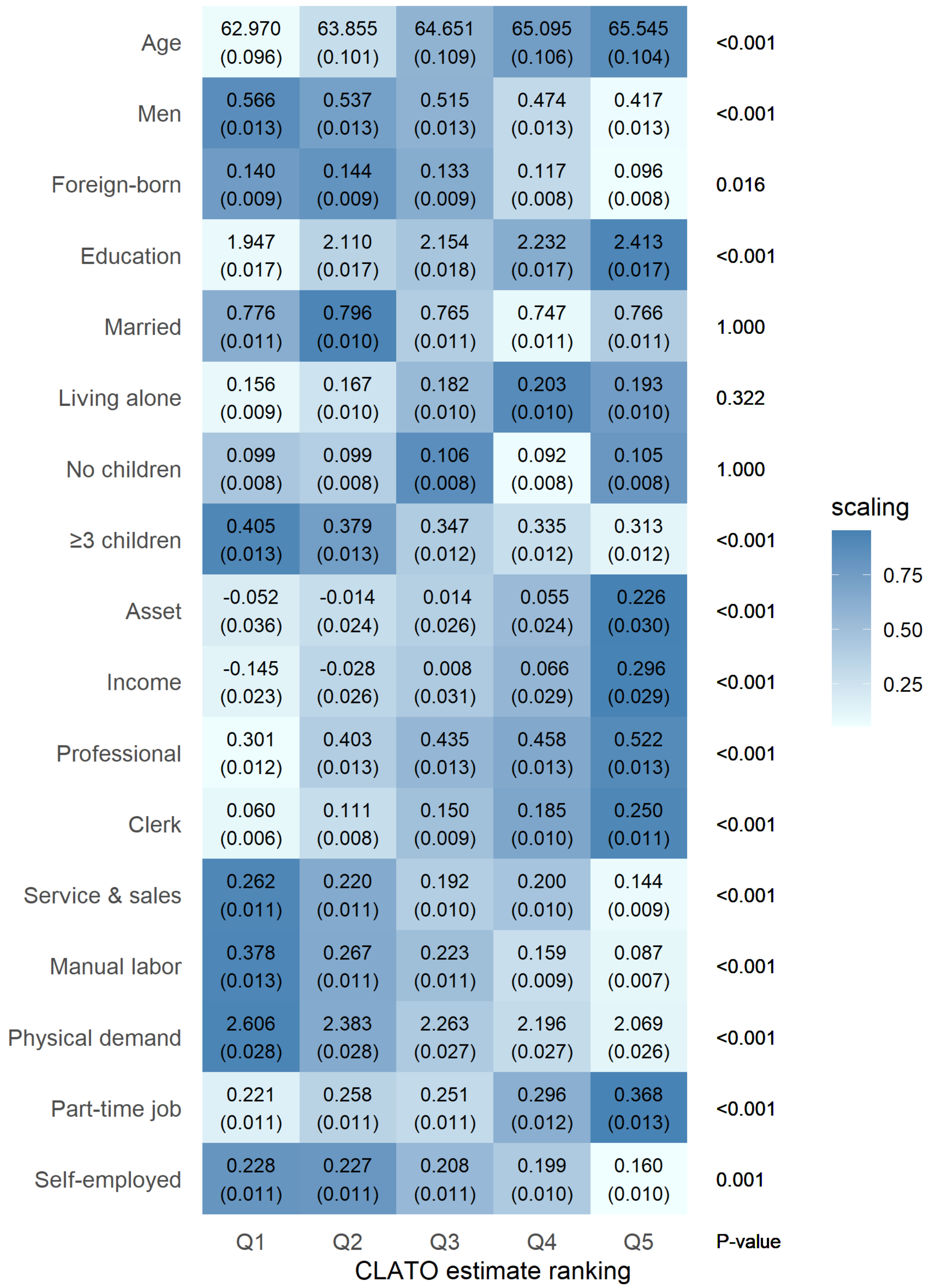
Figure 3.5 shows the heterogeneity of individual characteristics. Individuals in the highest CLATO group (Q5) tended to be older, female, and born in the country and had fewer than three children, higher education, assets, and income than those in the lowest group (Q1). Those who worked as clerks and whose jobs were professional or part-time tended to be categorized into the highest group, whereas those who worked in service and sales and manual jobs and whose jobs were physically demanding or self-employed tended to be categorized into the lowest group.

Figure 3.6 suggests that those with better health and well-being received more benefits from retirement. Specifically, those in the highest CLATO group tended to have better self-rated health, life satisfaction, and eyesight and hearing abilities than those in the lowest CLATO group. In contrast, those in the lowest group were more likely to have depressive symptoms, hypertension, diabetes, lung and heart diseases, arthritis, psychiatric problems, hyperlipidemia, health limitations in working, difficulties in activities of daily living and instrumental activities of daily living, and pain-related problems. Furthermore, while obese individuals tended to be categorized into the lowest group, those who frequently engaged in physical activity tended to be in the highest group.

Figure 3.7 shows the heterogeneity across countries. People in Denmark and Greece tended to fall into higher CLATO groups, whereas those in Estonia and France tended to fall into lower CLATO groups.

Appendix Figure H.4 presents the partial dependence plots for the continuous variables. The estimates of CLATO tended to increase with age until age 65 and then flattened after age 65. Those with average or below-average assets and income showed large variations in CLATO, whereas those with higher assets and income tended to have higher CLATO.

Figure 3.5 Heterogeneity in Individual Characteristics



Asset and income are standardized to z-scores. Q1 is the group with the lowest CLATO, while Q5 is the group with the highest CLATO. Each tile indicates the mean value of a covariate within the group and its standard deviation in parentheses. P-values of F-statistics are corrected using the Bonferroni method.

Figure 3.6 Heterogeneity in Health and Behaviors

	Q1	Q2	Q3	Q4	Q5	P-value
Baseline cognition	11.022 (0.085)	11.297 (0.081)	11.401 (0.083)	11.329 (0.081)	11.101 (0.085)	0.300
Self-rated health	2.800 (0.021)	3.138 (0.023)	3.353 (0.024)	3.674 (0.024)	3.986 (0.022)	<0.001
Depression	0.308 (0.030)	0.040 (0.027)	-0.001 (0.026)	-0.128 (0.024)	-0.290 (0.019)	<0.001
Life satisfaction	-0.292 (0.029)	-0.027 (0.026)	-0.004 (0.025)	0.129 (0.024)	0.329 (0.021)	<0.001
Hypertension	0.552 (0.013)	0.478 (0.013)	0.428 (0.013)	0.339 (0.012)	0.252 (0.011)	<0.001
Diabetes	0.188 (0.010)	0.155 (0.009)	0.114 (0.008)	0.107 (0.008)	0.067 (0.006)	<0.001
Cancer	0.067 (0.006)	0.080 (0.007)	0.067 (0.006)	0.077 (0.007)	0.079 (0.007)	1.000
Lung disease	0.066 (0.006)	0.046 (0.005)	0.051 (0.006)	0.040 (0.005)	0.034 (0.005)	0.029
Heart disease	0.155 (0.009)	0.115 (0.008)	0.128 (0.009)	0.093 (0.008)	0.071 (0.007)	<0.001
Stroke	0.034 (0.005)	0.027 (0.004)	0.026 (0.004)	0.017 (0.003)	0.015 (0.003)	0.169
Arthritis	0.478 (0.013)	0.377 (0.013)	0.312 (0.012)	0.250 (0.011)	0.219 (0.011)	<0.001
Psychiatric problems	0.144 (0.009)	0.118 (0.008)	0.108 (0.008)	0.088 (0.007)	0.072 (0.007)	<0.001
Hyperlipemia	0.462 (0.013)	0.345 (0.012)	0.297 (0.012)	0.231 (0.011)	0.123 (0.009)	<0.001
Health limitation	0.198 (0.010)	0.136 (0.009)	0.100 (0.008)	0.067 (0.006)	0.048 (0.006)	<0.001
Difficulty in ADL	0.090 (0.007)	0.053 (0.006)	0.041 (0.005)	0.029 (0.004)	0.017 (0.003)	<0.001
Difficulty in IADL	0.050 (0.006)	0.027 (0.004)	0.020 (0.004)	0.011 (0.003)	0.009 (0.002)	<0.001
Distance eyesight	3.574 (0.026)	3.675 (0.024)	3.753 (0.024)	3.849 (0.023)	3.954 (0.022)	<0.001
Near eyesight	3.327 (0.027)	3.479 (0.025)	3.631 (0.025)	3.758 (0.024)	3.860 (0.023)	<0.001
Hearing	3.026 (0.023)	3.356 (0.025)	3.573 (0.025)	3.778 (0.024)	4.026 (0.023)	<0.001
Pain problems	0.410 (0.013)	0.378 (0.013)	0.350 (0.012)	0.280 (0.012)	0.254 (0.011)	<0.001
Obesity	0.351 (0.012)	0.307 (0.012)	0.300 (0.012)	0.251 (0.011)	0.221 (0.011)	<0.001
Physical activity	0.844 (0.009)	0.869 (0.009)	0.876 (0.009)	0.889 (0.008)	0.907 (0.008)	<0.001
Heavy drinking	0.108 (0.008)	0.109 (0.008)	0.101 (0.008)	0.090 (0.007)	0.103 (0.008)	1.000
Smoking	0.162 (0.010)	0.157 (0.009)	0.162 (0.010)	0.155 (0.009)	0.146 (0.009)	1.000

CLATO estimate ranking

scaling: 0.75, 0.50, 0.25

Depression and life satisfaction are standardized to z-scores. Q1 is the group with the lowest CLATO, while Q5 is the group with the highest CLATO. Each tile indicates the mean value of a covariate within the group and its standard deviation in parentheses. P-values of F-statistics are corrected using the Bonferroni method.

Figure 3.7 Heterogeneity in Countries

	Q1	Q2	Q3	Q4	Q5	P-value
Austria	0.013 (0.003)	0.013 (0.003)	0.009 (0.002)	0.017 (0.003)	0.006 (0.002)	1.000
Belgium	0.021 (0.004)	0.022 (0.004)	0.025 (0.004)	0.020 (0.004)	0.018 (0.003)	1.000
Croatia	0.006 (0.002)	0.008 (0.002)	0.008 (0.002)	0.004 (0.002)	0.008 (0.002)	1.000
Czech Republic	0.049 (0.006)	0.043 (0.005)	0.038 (0.005)	0.031 (0.005)	0.022 (0.004)	0.060
Denmark	0.020 (0.004)	0.029 (0.004)	0.036 (0.005)	0.060 (0.006)	0.113 (0.008)	<0.001
Estonia	0.077 (0.007)	0.077 (0.007)	0.076 (0.007)	0.061 (0.006)	0.026 (0.004)	<0.001
France	0.051 (0.006)	0.027 (0.004)	0.024 (0.004)	0.020 (0.004)	0.018 (0.003)	<0.001
Germany	0.047 (0.005)	0.064 (0.006)	0.052 (0.006)	0.048 (0.006)	0.051 (0.006)	1.000
Greece	0.016 (0.003)	0.019 (0.004)	0.035 (0.005)	0.043 (0.005)	0.043 (0.005)	<0.001
Israel	0.007 (0.002)	0.014 (0.003)	0.021 (0.004)	0.022 (0.004)	0.023 (0.004)	0.113
Italy	0.025 (0.004)	0.016 (0.003)	0.024 (0.004)	0.018 (0.003)	0.024 (0.004)	1.000
Luxembourg	0.007 (0.002)	0.009 (0.003)	0.010 (0.003)	0.011 (0.003)	0.010 (0.003)	1.000
Poland	0.011 (0.003)	0.008 (0.002)	0.002 (0.001)	0.005 (0.002)	0.003 (0.001)	0.796
Slovenia	0.029 (0.004)	0.026 (0.004)	0.015 (0.003)	0.016 (0.003)	0.018 (0.003)	0.968
Spain	0.021 (0.004)	0.026 (0.004)	0.026 (0.004)	0.025 (0.004)	0.016 (0.003)	1.000
Sweden	0.060 (0.006)	0.056 (0.006)	0.052 (0.006)	0.059 (0.006)	0.052 (0.006)	1.000
Switzerland	0.044 (0.005)	0.047 (0.006)	0.057 (0.006)	0.051 (0.006)	0.063 (0.006)	1.000
England	0.189 (0.010)	0.191 (0.010)	0.186 (0.010)	0.184 (0.010)	0.176 (0.010)	1.000
United States	0.305 (0.012)	0.304 (0.012)	0.304 (0.012)	0.305 (0.012)	0.308 (0.012)	1.000

CLATO estimate ranking

scaling: 0.75, 0.50, 0.25

Q1 is the group with the lowest CLATO, while Q5 is the group with the highest CLATO. Each tile indicates the mean value of a covariate within the group and its standard deviation in parentheses. P-values of F-statistics are corrected using the Bonferroni method.

3.4.4 Monetary Cost Estimation of Increasing ORA

Based on the LATO estimates for the quintile groups shown in Figure 3.4, we estimate that the total monetary cost of dementia care will increase by 5.0 billion dollars (1.4%) in the United States and 3.3 billion pounds (5.2%) in the United Kingdom in 2030 due to an increase in their ORA from age 65 to 66 (See Appendix Tables H.3 and H.4 for details). In the United States, the impact of increasing the ORA will be limited because a large proportion of workers retire at the ERA of 62 years. In contrast, the impact of increasing the ORA in the United Kingdom will not be negligible because most workers consider retirement at the ORA, given that its pension system has no early retirement scheme. Our findings suggest that delayed retirement owing to increased ORA use increases the monetary costs of dementia care.

3.4.5 Robustness Checks

Table 3.4 displays the results of the robustness checks comparing the LATO estimated using different models. Column (1) indicates the estimates from a model that restricts participants to individuals aged 55–75 years. Column (2) shows a model that excludes those who mentioned retirement but worked in the second wave (i.e., partly retired) and examines the impact of full retirement on cognitive function. Column (3) shows a model that excludes those who engaged in part-time jobs or were self-employed in the first wave and studied only full-time employees. Column (4) shows a model that excludes data from the United States, given that it has the largest sample size in our dataset. As shown in Table 3.4, all the LATO estimates are similar to our main results, which suggests that our findings are robust, even in different settings.

Table 3.4 Robustness Checks

	(1)	(2)	(3)	(4)
Retirement	1.377***	1.334**	1.366*	1.348**
	(0.528)	(0.578)	(0.718)	(0.599)
Observations	7268	6128	4582	5218

Column (1) is for a model that restricted participants to individuals aged from 55 to 75 years. Column (2) is for a model that excluded those who mentioned retirement but worked in the second wave. Column (3) is for a model that excluded those who engaged in a part-time job or were self-employed in the first wave. Column (4) is for a model that excluded data from the United States.
Standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.5 Discussion and Conclusions

This study investigated the heterogeneous treatment effects of retirement on cognitive function using data from 19 countries. We found that retirees had better cognitive function than workers on average and that the conditional average treatment effects varied depending on the individual's characteristics.

Our findings on ATE were consistent with those of previous studies suggesting that retirement improves cognitive function (Bianchini & Borella, 2016; de Grip et al., 2015). We found that the estimates of the OLS and non-IV forests indicated non-significant associations, but they could be negatively biased because of health selection for retirement. After eliminating potential endogeneity using IV methods, both the conventional 2SLS and IV forests showed beneficial associations with retirement and cognitive function. The estimate of 1.348 words in the IV forests is large, given that it corresponds to a 0.42 standard deviation of the distribution of the cognitive function score.[35] Improvements in cognitive function after retirement can be explained through several pathways. First, job strain is a potential risk factor for decreased cognitive function (Agbenyikey et al., 2015; Elovainio et al., 2009), but retirement releases individuals from psychosocial stress. Second, retirees can invest more time in their health than workers. Many studies have shown that retirement is associated with healthy behaviors such as increased physical activity, sleep quality, and smoking cessation (Celidoni & Rebba, 2017; Eibich, 2015; Insler, 2014; Kämpfen & Maurer, 2016; Kesavayuth et al., 2018; Müller & Shaikh, 2018; Myllyntausta et al., 2018; Syse et al., 2017), which protect against cognitive decline.

This study presented the heterogeneous treatment effect of retirement on cognitive function depending on individuals' characteristics. For example, we found that women tended to have higher CLATO. This gender difference is in line with the evidence that women are more likely to engage in exercise to maintain their physical and mental health after retirement than men, which could induce disparities in cognitive function (Atalay et al., 2019; Sato & Noguchi, 2023). Additionally, individuals who engaged in physical activity before retirement tended to have a higher CLATO. We assumed that they had a habit of exercising and maintained it after retirement, which was beneficial to their cognitive function. Thus, some heterogeneity may be explained by post-retirement health behaviors.

[35] Kraft (2020) reviewed 747 randomized control trials for cognitive interventions and proposed that the effect size of over 0.2 standard deviation can be interpreted as large.

Furthermore, our findings regarding heterogeneity are consistent with those of the Grossman model (Grossman, 1972). We found that people with higher levels of education, assets, and income tended to receive more benefits from retirement. According to this model, retirees have more time, but a lower budget, to invest in their health. However, people with a high socioeconomic status can afford health investments to improve their cognitive function. The tendency for people with better health before retirement to have higher cognitive function after retirement was also consistent with the model because healthy people have more time to spend on health investments than sick people. The association between health limitations before retirement and cognitive function after retirement is consistent with the findings of an empirical study (Denier et al., 2017).

Regarding the characteristics of the pre-retirement job in relation to cognitive decline, people who retired from a professional occupation tended to have a higher CLATO, whereas those who retired from manual labor and physically demanding jobs tended to have a lower CLATO. This finding was consistent with many studies showing an association between retirement from highly complex and mentally demanding jobs and a slower rate of cognitive decline (Carr et al., 2020; Fisher et al., 2014; Kajitani et al., 2017; Vélez-Coto et al., 2021). In contrast, we also found that retirement from physically demanding jobs was less beneficial for cognitive function. Similarly, previous studies have shown that engagement in physically demanding jobs was associated with a higher risk of dementia (Nabe-Nielsen et al., 2021; Smyth et al., 2004; Zotcheva et al., 2023). Occupations with high physical demands are often combined with low job control, which is less cognitively stimulating (Andel et al., 2015; Dong et al., 2018; Romero Starke et al., 2019).These findings suggest that stimulation of cognitive function at the workplace gets carried over into retirement.

This study had several limitations. First, we only investigated the short-term effects of retirement, given that cognitive function was measured two years after retirement. Further studies are required to examine the long-term effects. Second, our analysis of heterogeneity based on specific covariates was exploratory. For example, we found that people with higher education and a professional job tended to have higher CLATO, but these characteristics could be confounded (i.e., people with higher education were likely to have a professional job). Therefore, confirmatory studies are necessary to determine the causal heterogeneity of specific covariates. Third, we could not capture the heterogeneity stemming from unmeasured covariates, although we included 60 candidate variables. Other factors, such as traumatic brain injury, social isolation, and air pollution, may modify the effect of retirement on cognitive function (Livingston et al., 2020). Fourth, the measured variables may have been subject to measurement errors because they were collected through self-reported interviews. However, it has been shown that the word recall test can predict the onset of dementia (Tierney et al., 2010). Additionally, asking for self-recognition of labor force status is meaningful because it can influence individuals' behavior (Eibich, 2015). Fifth, the generalizability of our findings may be limited to Western countries. Given that Asian countries face more rapid population aging than Western countries and are also increasing their SPA, analyses using harmonized data such as the China Health and Retirement Longitudinal Study and the Korean Longitudinal Study of Aging will provide essential and comparable evidence for these countries. Sixth, although data harmonization was performed by field experts (Angrisani & Lee, 2012a, 2012b; Hu & Lee, 2012; Shih et al., 2012; Zamarro & Lee, 2012), discrepancies across surveys may remain. However, our pooling analysis provided important insights into the heterogeneity across countries.

In summary, we found that the impact of retirement on cognitive function varied depending on the individual's characteristics. Therefore, we recommend that policymakers provide options for early retirement in the pension system to allow individuals to decide when to retire. If a prediction tool developed using this machine-learning model is launched, individuals will have the capability to determine the optimal timing for their retirement by inputting their characteristics into the program. Given that retirement improves cognitive function, the balance between the social benefits of increasing the state pension age and the individual cost of increasing dementia risk due to delayed retirement should be considered.

References

Agahi, N., Fors, S., Fritzell, J., & Shaw, B. A. (2018). Smoking and Physical Inactivity as Predictors of Mobility Impairment During Late Life: Exploring Differential Vulnerability Across Education Level in Sweden. *The Journals of Gerontology: Series B*, *73*(4), 675–683. https://doi.org/10.1093/geronb/gbw090

Agbenyikey, W., Karasek, R., Cifuentes, M., Wolf, P. A., Seshadri, S., Taylor, J. A., Beiser, A. S., & Au, R. (2015). Job strain and cognitive decline: A prospective study of the framingham offspring cohort. *The International Journal of Occupational and Environmental Medicine*, *6*(2), 79–94. https://doi.org/10.15171/ijoem.2015.534

Andel, R., Infurna, F. J., Hahn Rickenbach, E. A., Crowe, M., Marchiondo, L., & Fisher, G. G. (2015). Job strain and trajectories of change in episodic memory before and after retirement: Results from the Health and Retirement Study. *Journal of Epidemiology and Community Health*, *69*(5), 442–446. https://doi.org/10.1136/jech-2014-204754

Andersen, A. J., Hecker, I., Wallez, S., Witteveen, A., Lora, A., Mittendorfer-Rutz, E., Corrao, G., Walter, H., Haro, J. M., Sijbrandij, M., Compagnoni, M. M., Felez-Nobrega, M., Kalisch, R., Bryant, R., Melchior, M., & Mary-Krause, M. (2023). Are we equally at risk of changing smoking behavior during a public health crisis? Impact of educational level on smoking from the TEMPO cohort. *BMC Public Health*, *23*, 1016. https://doi.org/10.1186/s12889-023-15799-1

Angrisani, M., & Lee, J. (2012a). *Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Income Measures*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR861z5.html

Angrisani, M., & Lee, J. (2012b). *Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Wealth Measures*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR861z6.html

Arrighi, H. M., & Hertz-Picciotto, I. (1994). The evolving concept of the healthy worker survivor effect. *Epidemiology*, *5*(2), 189–196. https://doi.org/10.1097/00001648-199403000-00009

Atalay, K., Barrett, G. F., & Staneva, A. (2019). The effect of retirement on elderly cognitive functioning. *Journal of Health Economics*, *66*, 37–53. https://doi.org/10.1016/j.jhealeco.2019.04.006

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7353–7360. https://doi.org/10.1073/pnas.1510489113

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, *47*(2), 1148–1178. https://doi.org/10.1214/18-AOS1709

Bale, T. L., & Epperson, C. N. (2015). Sex differences and stress across the lifespan. *Nature Neuroscience*, *18*(10), Article 10. https://doi.org/10.1038/nn.4112

Barnett, I., van Sluijs, E. M. F., & Ogilvie, D. (2012). Physical activity and transitioning to retirement: A systematic review. *American Journal of Preventive Medicine*, *43*(3), 329–336. https://doi.org/10.1016/j.amepre.2012.05.026

Behncke, S. (2012). Does retirement trigger ill health? *Health Economics*, *21*(3), 282–300. https://doi.org/10.1002/hec.1712

Bianchini, L., & Borella, M. (2016). Retirement and memory in Europe. *Ageing & Society*, *36*, 1434–1458. https://doi.org/10.1017/S0144686X15000434

Biewen, M., & Kugler, P. (2021). Two-stage least squares random forests with an application to Angrist and Evans (1998). *Economics Letters*, *204*, 109893. https://doi.org/10.1016/j.econlet.2021.109893

Bingley, P., & Martinello, A. (2013). Mental retirement and schooling. *European Economic Review*, *63*, 292–298. https://doi.org/10.1016/j.euroecorev.2013.01.004

Blake, H., & Garrouste, C. (2019). Collateral Effects of a Pension Reform in France. *Annals of Economics and Statistics*, *133*, 57–86. https://doi.org/10.15609/annaeconstat2009.133.0057

Bogaard, L. V. D., Henkens, K., & Kalmijn, M. (2014). So now what? Effects of retirement on civic engagement. *Ageing & Society*, *34*(7), 1170–1192. https://doi.org/10.1017/S0144686X13000019

Bond, J., Dickinson, H. O., Matthews, F., Jagger, C., Brayne, C., & MRC CFAS. (2006). Self-rated health status as a predictor of death, functional and cognitive impairment: A longitudinal cohort study. *European Journal of Ageing*, *3*(4), 193–206. https://doi.org/10.1007/s10433-006-0039-8

Bonsang, E., Adam, S., & Perelman, S. (2012). Does retirement affect cognitive functioning? *Journal of Health Economics*, *31*(3), 490–501. https://doi.org/10.1016/j.jhealeco.2012.03.005

Brenowitz, W. D., Hubbard, R. A., Crane, P. K., Gray, S. L., Zaslavsky, O., & Larson, E. B. (2014). Longitudinal Associations between Self-Rated Health and Performance-Based

Physical Function in a Population-Based Cohort of Older Adults. *PLOS ONE*, *9*(11), e111761. https://doi.org/10.1371/journal.pone.0111761

Calvo, E., Sarkisian, N., & Tamborini, C. R. (2013). Causal Effects of Retirement Timing on Subjective Physical and Emotional Health. *The Journals of Gerontology: Series B*, *68*(1), 73–84. https://doi.org/10.1093/geronb/gbs097

Carr, D. C., Willis, R., Kail, B. L., & Carstensen, L. L. (2020). Alternative Retirement Paths and Cognitive Performance: Exploring the Role of Preretirement Job Complexity. *The Gerontologist*, *60*(3), 460–471. https://doi.org/10.1093/geront/gnz079

Celidoni, M., Dal Bianco, C., & Weber, G. (2017). Retirement and cognitive decline. A longitudinal analysis using SHARE data. *Journal of Health Economics*, *56*, 113–125. https://doi.org/10.1016/j.jhealeco.2017.09.003

Celidoni, M., & Rebba, V. (2017). Healthier lifestyles after retirement in Europe? Evidence from SHARE. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care*, *18*(7), 805–830. https://doi.org/10.1007/s10198-016-0828-8

Centers for Disease Control and Prevention. (2022, January 6). *Binge Drinking*. Centers for Disease Control and Prevention. https://www.cdc.gov/alcohol/fact-sheets/binge-drinking.htm

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1–C68. https://doi.org/10.1111/ectj.12097

Chung, S., Domino, M. E., & Stearns, S. C. (2009). The effect of retirement on weight. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *64*(5), 656–665. https://doi.org/10.1093/geronb/gbn044

Cimler, R., Maresova, P., Kuhnova, J., & Kuca, K. (2019). Predictions of Alzheimer's disease treatment and care costs in European countries. *PloS One*, *14*(1), e0210958. https://doi.org/10.1371/journal.pone.0210958

Coe, N. B., & Lindeboom, M. (2008). Does Retirement Kill You? Evidence from Early Retirement Windows. *IZA Discussion Paper*, *No. 3817*. https://doi.org/10.2139/ssrn.1298260

Coe, N. B., von Gaudecker, H.-M., Lindeboom, M., & Maurer, J. (2012). The effect of retirement on cognitive functioning. *Health Economics*, *21*(8), 913–927. https://doi.org/10.1002/hec.1771

Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, *30*(1), 77–86. https://doi.org/10.1016/j.jhealeco.2010.11.002

Comi, S. L., Cottini, E., & Lucifora, C. (2022). The effect of retirement on social relationships. *German Economic Review*, *23*(2), 275–299. https://doi.org/10.1515/ger-2020-0109

Correia, S. (2015, November). *Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix*. http://scorreia.com/research/singletons.pdf

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, *96*(1), 187–199. https://doi.org/10.1093/biomet/asn055

Dave, D., Rashad, I., & Spasojevic, J. (2008). The Effects of Retirement on Physical and Mental Health Outcomes. *Southern Economic Journal*, *75*(2), 497–523. https://doi.org/10.1002/j.2325-8012.2008.tb00916.x

de Grip, A., Dupuy, A., Jolles, J., & van Boxtel, M. (2015). Retirement and cognitive development in the Netherlands: Are the retired really inactive? *Economics and Human Biology*, *19*, 157–169. https://doi.org/10.1016/j.ehb.2015.08.004

Denier, N., Clouston, S. A. P., Richards, M., & Hofer, S. M. (2017). Retirement and Cognition: A Life Course View. *Advances in Life Course Research*, *31*, 11–21. https://doi.org/10.1016/j.alcr.2016.10.004

Dong, L., Eaton, W. W., Spira, A. P., Agnew, J., Surkan, P. J., & Mojtabai, R. (2018). Job strain and cognitive change: The Baltimore Epidemiologic Catchment Area follow-up study. *Occupational and Environmental Medicine*, *75*(12), 856–862. https://doi.org/10.1136/oemed-2018-105213

Ebeid, M., & Oguzoglu, U. (2023). Short-term effect of retirement on health: Evidence from nonparametric fuzzy regression discontinuity design. *Health Economics*. https://doi.org/10.1002/hec.4669

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, *43*, 1–12. https://doi.org/10.1016/j.jhealeco.2015.05.001

Elovainio, M., Ferrie, J. E., Singh-Manoux, A., Gimeno, D., De Vogli, R., Shipley, M. J., Vahtera, J., Brunner, E. J., Marmot, M. G., & Kivimäki, M. (2009). Cumulative exposure to high-strain and active jobs as predictors of cognitive function: The Whitehall II study. *Occupational and Environmental Medicine*, *66*(1), 32–37. https://doi.org/10.1136/oem.2008.039305

Fisher, G. G., Stachowski, A., Infurna, F. J., Faul, J. D., Grosch, J., & Tetrick, L. E. (2014). Mental work demands, retirement, and longitudinal trajectories of cognitive functioning. *Journal of Occupational Health Psychology*, *19*(2), 231–242. https://doi.org/10.1037/a0035724

Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., Pletcher, M. A., Smith, A. E., Tang, K., Yuan, C.-W., Brown, J. C., Friedman, J., He, J., Heuton, K. R., Holmberg, M., Patel, D. J., Reidy, P., Carter, A., Cercy, K., … Murray, C. J. L. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet (London, England)*, *392*(10159), 2052–2090. https://doi.org/10.1016/S0140-6736(18)31694-5

Garrouste, C., & Perdrix, E. (2022). Is there a consensus on the health consequences of retirement? A literature review. *Journal of Economic Surveys*, *36*(4), 841–879. https://doi.org/10.1111/joes.12466

GBD 2019 Diseases and Injuries Collaborators. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet (London, England)*, *396*(10258), 1204–1222. https://doi.org/10.1016/S0140-6736(20)30925-9

Glymour, M. M., & Avendano, M. (2009). Can self-reported strokes be used to study stroke incidence and risk factors?: Evidence from the health and retirement study. *Stroke*, *40*(3), 873–879. https://doi.org/10.1161/STROKEAHA.108.529479

Glymour, M. M., & Swanson, S. A. (2020). Instrumental Variables and Quasi-Experimental Approaches. In T. L. Lash, T. J. VanderWeele, S. Haneuse, & K. J. Rothman (Eds.), *Modern Epidemiology* (4th edition, pp. 677–709). Wolters Kluwer.

Gorry, A., Gorry, D., & Slavov, S. N. (2018). Does retirement improve health and life satisfaction? *Health Economics*, *27*(12), 2067–2086. https://doi.org/10.1002/hec.3821

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, *80*(2), 223–255.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, *50*(4), 1029–1054. https://doi.org/10.2307/1912775

Helldán, A., Lallukka, T., Rahkonen, O., & Lahelma, E. (2012). Changes in healthy food habits after transition to old age retirement. *European Journal of Public Health*, *22*(4), 582–586. https://doi.org/10.1093/eurpub/ckr060

Hessel, P. (2016). Does retirement (really) lead to worse health among European men and women across all educational levels? *Social Science & Medicine (1982)*, *151*, 19–26. https://doi.org/10.1016/j.socscimed.2015.12.018

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Honaker, J., & King, G. (2010). What to do About Missing Values in Time Series Cross-Section Data. *American Journal of Political Science*, *54*(3), 561–581.

Hu, P., & Lee, J. (2012). *Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Chronic Medical Conditions*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR861z1.html

Hurd, M. D., Martorell, P., Delavande, A., Mullen, K. J., & Langa, K. M. (2013). Monetary Costs of Dementia in the United States. *The New England Journal of Medicine*, *368*(14), 1326–1334. https://doi.org/10.1056/NEJMsa1204629

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, *38*(1), 21–37.

Insler, M. (2014). The Health Consequences of Retirement. *Journal of Human Resources*, *49*(1), 195–233. https://doi.org/10.3368/jhr.49.1.195

Jia, J., Zhao, T., Liu, Z., Liang, Y., Li, F., Li, Y., Liu, W., Li, F., Shi, S., Zhou, C., Yang, H., Liao, Z., Li, Y., Zhao, H., Zhang, J., Zhang, K., Kan, M., Yang, S., Li, H., … Cummings, J. (2023). Association between healthy lifestyle and memory decline in older adults: 10 year, population based, prospective cohort study. *BMJ*, *380*, e072691. https://doi.org/10.1136/bmj-2022-072691

Johnston, D. W., & Lee, W.-S. (2009). Retiring to the good life? The short-term effects of retirement on health. *Economics Letters*, *103*(1), 8–11. https://doi.org/10.1016/j.econlet.2009.01.015

Kajitani, S., Sakata, K., & Mckenzie, C. (2017). Occupation, retirement and cognitive functioning. *Ageing & Society*, *37*(8), 1568–1596. https://doi.org/10.1017/S0144686X16000465

Kämpfen, F., & Maurer, J. (2016). Time to burn (calories)? The impact of retirement on physical activity among mature Americans. *Journal of Health Economics*, *45*, 91–102. https://doi.org/10.1016/j.jhealeco.2015.12.001

Kesavayuth, D., Rosenman, R. E., & Zikos, V. (2018). Retirement and health behaviour. *Applied Economics*, *50*(54), 5859–5876. https://doi.org/10.1080/00036846.2018.1488070

Kesse-Guyot, E., Andreeva, V. A., Lassale, C., Hercberg, S., & Galan, P. (2014). Clustering of midlife lifestyle behaviors and subsequent cognitive function: A longitudinal study. *American Journal of Public Health*, *104*(11), e170-177. https://doi.org/10.2105/AJPH.2014.302121

Kikuchi, H., Inoue, S., Fukushima, N., Takamiya, T., Odagiri, Y., Ohya, Y., Amagasa, S., Oka, K., & Owen, N. (2017). Social participation among older adults not engaged in full- or part-time work is associated with more physical activity and less sedentary time. *Geriatrics & Gerontology International*, *17*(11), 1921–1927. https://doi.org/10.1111/ggi.12995

Kirkwood, T. B. L. (2008). A systematic look at an old problem. *Nature*, *451*(7179), Article 7179. https://doi.org/10.1038/451644a

Kivimäki, M., Nyberg, S. T., Batty, G. D., Fransson, E. I., Heikkilä, K., Alfredsson, L., Bjorner, J. B., Borritz, M., Burr, H., Casini, A., Clays, E., De Bacquer, D., Dragano, N., Ferrie, J. E., Geuskens, G. A., Goldberg, M., Hamer, M., Hooftman, W. E., Houtman, I. L., … Theorell, T. (2012). Job strain as a risk factor for coronary heart disease: A collaborative meta-analysis of individual participant data. *The Lancet*, *380*(9852), 1491–1497. https://doi.org/10.1016/S0140-6736(12)60994-5

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, *133*(1), 97–126. https://doi.org/10.1016/j.jeconom.2005.02.011

Kobayashi, E., Sugawara, I., Fukaya, T., Okamoto, S., & Liang, J. (2022). Retirement and Social Activities in Japan: Does Age Moderate the Association? *Research on Aging*, *44*(2), 144–155. https://doi.org/10.1177/01640275211005185

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Lee, J., Phillips, D., Wilkens, J., & Gateway to Global Aging Data Team. (2021). Gateway to Global Aging Data: Resources for Cross-National Comparisons of Family, Social Environment, and Healthy Aging. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *76*(Suppl 1), S5–S16. https://doi.org/10.1093/geronb/gbab050

Lei, X., & Liu, H. (2018). Gender difference in the impact of retirement on cognitive abilities: Evidence from urban China. *Journal of Comparative Economics*, *46*(4), 1425–1446. https://doi.org/10.1016/j.jce.2018.01.005

Lerner, D. J., Levine, S., Malspeis, S., & D'Agostino, R. B. (1994). Job strain and health-related quality of life in a national sample. *American Journal of Public Health*, *84*(10), 1580–1585. https://doi.org/10.2105/AJPH.84.10.1580

Li, Z. (2022). *The Social Security Retirement Age* (CRC Report R44670). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R44670

Lin, W., & Wooldridge, J. M. (2019). Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models. In M. Tsionas (Ed.), *Panel Data Econometrics* (pp. 21–43). Academic Press. https://doi.org/10.1016/B978-0-12-814367-4.00002-2

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., … Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Maurage, P., Joassin, F., Speth, A., Modave, J., Philippot, P., & Campanella, S. (2012). Cerebral effects of binge drinking: Respective influences of global alcohol intake and consumption pattern. *Clinical Neurophysiology*, *123*(5), 892–901. https://doi.org/10.1016/j.clinph.2011.09.018

Mayer, M. (2021). *missRanger: Fast Imputation of Missing Values* (2.1.3) [Computer software]. https://cran.r-project.org/web/packages/missRanger/missRanger.pdf

Mazzonna, F., & Peracchi, F. (2012). Ageing, cognitive abilities and retirement. *European Economic Review*, *56*(4), 691–710. https://doi.org/10.1016/j.euroecorev.2012.03.004

Mazzonna, F., & Peracchi, F. (2017). Unhealthy Retirement? *Journal of Human Resources*, *52*(1), 128–151. https://doi.org/10.3368/jhr.52.1.0914-6627R1

Messe, P.-J., & Wolff, F.-C. (2019). The short-term effects of retirement on health within couples: Evidence from France. *Social Science & Medicine (1982)*, *221*, 27–39. https://doi.org/10.1016/j.socscimed.2018.12.008

Moon, J. R., Glymour, M. M., Subramanian, S. V., Avendaño, M., & Kawachi, I. (2012). Transition to retirement and risk of cardiovascular disease: Prospective analysis of the US health and retirement study. *Social Science & Medicine*, *75*(3), 526–530. https://doi.org/10.1016/j.socscimed.2012.04.004

Moore, A. A., Endo, J. O., & Carter, M. K. (2003). Is There a Relationship Between Excessive Drinking and Functional Impairment in Older Persons? *Journal of the*

American Geriatrics Society, *51*(1), 44–49. https://doi.org/10.1034/j.1601-5215.2002.51008.x

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165. https://doi.org/10.1212/wnl.39.9.1159

Müller, T., & Shaikh, M. (2018). Your retirement and my health behavior: Evidence on retirement externalities from a fuzzy regression discontinuity design. *Journal of Health Economics*, *57*, 45–59. https://doi.org/10.1016/j.jhealeco.2017.10.005

Myllyntausta, S., Salo, P., Kronholm, E., Pentti, J., Kivimäki, M., Vahtera, J., & Stenholm, S. (2018). Changes in Sleep Difficulties During the Transition to Statutory Retirement. *Sleep*, *41*(1), zsx182. https://doi.org/10.1093/sleep/zsx182

Nabe-Nielsen, K., Holtermann, A., Gyntelberg, F., Garde, A. H., Islamoska, S., Prescott, E., Schnohr, P., & Hansen, Å. M. (2021). The effect of occupational physical activity on dementia: Results from the Copenhagen Male Study. *Scandinavian Journal of Medicine & Science in Sports*, *31*(2), 446–455. https://doi.org/10.1111/sms.13846

Narayanan, L., Menon, S., & Spector, P. E. (1999). Stress in the workplace: A comparison of gender and occupations. *Journal of Organizational Behavior*, *20*(1), 63–73. https://doi.org/10.1002/(SICI)1099-1379(199901)20:1<63::AID-JOB873>3.0.CO;2-J

National Institute on Alcohol Abuse and Alcoholism. (2023, April 10). *The Healthcare Professional's Core Resource on Alcohol*. National Institute on Alcohol Abuse and Alcoholism. https://www.niaaa.nih.gov/health-professionals-communities/core-resource-on-alcohol/basics-defining-how-much-alcohol-too-much#pub-toc4

Neuman, K. (2008). Quit Your Job and Get Healthier? The Effect of Retirement on Health. *Journal of Labor Research*, *29*(2), 177–201. https://doi.org/10.1007/s12122-007-9036-8

Nishimura, Y., Oikawa, M., & Motegi, H. (2018). What Explains the Difference in the Effect of Retirement on Health? Evidence from Global Aging Data. *Journal of Economic Surveys*, *32*(3), 792–847. https://doi.org/10.1111/joes.12215

Okusaga, O., Stewart, M. C. W., Butcher, I., Deary, I., Fowkes, F. G. R., & Price, J. F. (2013). Smoking, hypercholesterolaemia and hypertension as risk factors for cognitive impairment in older adults. *Age and Ageing*, *42*(3), 306–311. https://doi.org/10.1093/ageing/afs193

Organisation for Economic Co-operation and Development. (2021). *Pensions at a Glance 2021: OECD and G20 Indicators*. OECD Publishing. https://doi.org/10.1787/ca401ebd-en

Organisation for Economic Co-operation and Development & World Health Organization. (2020). *Health at a Glance: Asia/Pacific 2020: Measuring Progress Towards Universal Health Coverage*. OECD Publishing. https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-asia-pacific-2020_26b007cd-en

Regitz-Zagrosek, V., & Kararigas, G. (2017). Mechanistic Pathways of Sex Differences in Cardiovascular Disease. *Physiological Reviews*, *97*(1), 1–37. https://doi.org/10.1152/physrev.00021.2015

Robinson, P. (1988). Root- N-Consistent Semiparametric Regression. *Econometrica*, *56*(4), 931–954.

Roehrig, B., Hoeffken, K., Pientka, L., & Wedding, U. (2007). How many and which items of activities of daily living (ADL) and instrumental activities of daily living (IADL) are necessary for screening. *Critical Reviews in Oncology/Hematology*, *62*(2), 164–171. https://doi.org/10.1016/j.critrevonc.2006.10.001

Romero Starke, K., Seidler, A., Hegewald, J., Klimova, A., & Palmer, K. (2019). Retirement and decline in episodic memory: Analysis from a prospective study of adults in England. *International Journal of Epidemiology*, *48*(6), 1925–1936. https://doi.org/10.1093/ije/dyz135

Rose, L. (2020). Retirement and health: Evidence from England. *Journal of Health Economics*, *73*, 102352. https://doi.org/10.1016/j.jhealeco.2020.102352

Sabia, S., Nabi, H., Kivimaki, M., Shipley, M. J., Marmot, M. G., & Singh-Manoux, A. (2009). Health behaviors from early to late midlife as predictors of cognitive function: The Whitehall II study. *American Journal of Epidemiology*, *170*(4), 428–437. https://doi.org/10.1093/aje/kwp161

Salthouse, T. A. (2006). Mental Exercise and Mental Aging: Evaluating the Validity of the "Use It or Lose It" Hypothesis. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *1*(1), 68–87. https://doi.org/10.1111/j.1745-6916.2006.00005.x

Sato, K., Kondo, N., Hanazato, M., Tsuji, T., & Kondo, K. (2021). Potential causal effect of physical activity on reducing the risk of dementia: A 6-year cohort study from the Japan Gerontological Evaluation Study. *The International Journal of Behavioral*

Nutrition and Physical Activity, *18*(1), 140. https://doi.org/10.1186/s12966-021-01212-w

Sato, K., & Noguchi, H. (2023). *Sex Differences in the Impact of Retirement on Health: Evidence from 35 Countries* (SSRN Scholarly Paper 4413645). https://doi.org/10.2139/ssrn.4413645

Sato, K., Noguchi, H., & Inoue, K. (2023). *Heterogeneous Treatment Effect of Retirement on Cognitive Function* (SSRN Scholarly Paper 4577313). https://doi.org/10.2139/ssrn.4577313

Sato, K., Noguchi, H., Inoue, K., Kawachi, I., & Kondo, N. (2023). Retirement and cardiovascular disease: A longitudinal study in 35 countries. *International Journal of Epidemiology*, *52*(4), 1047–1059. https://doi.org/10.1093/ije/dyad058

Schmaus, B. J., Laubmeier, K. K., Boquiren, V. M., Herzer, M., & Zakowski, S. G. (2008). Gender and stress: Differential psychophysiological reactivity to stress reexposure in the laboratory. *International Journal of Psychophysiology*, *69*(2), 101–106. https://doi.org/10.1016/j.ijpsycho.2008.03.006

Shih, R. A., Lee, J., & Das, L. (2012). *Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Cognition*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR861z7.html

Smyth, K. A., Fritsch, T., Cook, T. B., McClendon, M. J., Santillan, C. E., & Friedland, R. P. (2004). Worker functions and traits associated with occupations and the development of AD. *Neurology*, *63*(3), 498–503. https://doi.org/10.1212/01.wnl.0000133007.87028.09

Ståhl, T., Rütten, A., Nutbeam, D., Bauman, A., Kannas, L., Abel, T., Lüschen, G., Rodriquez, D. J. A., Vinck, J., & van der Zee, J. (2001). The importance of the social environment for physically active lifestyle—Results from an international study. *Social Science & Medicine*, *52*(1), 1–10. https://doi.org/10.1016/S0277-9536(00)00116-7

Steinberg, M. L., Williams, J. M., & Li, Y. (2015). Poor Mental Health and Reduced Decline in Smoking Prevalence. *American Journal of Preventive Medicine*, *49*(3), 362–369. https://doi.org/10.1016/j.amepre.2015.01.016

Stenholm, S., Solovieva, S., Viikari-Juntura, E., Aalto, V., Kivimäki, M., & Vahtera, J. (2017). Change in body mass index during transition to statutory retirement: An occupational cohort study. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 85. https://doi.org/10.1186/s12966-017-0539-2

Stenholm, S., Westerlund, H., Salo, P., Hyde, M., Pentti, J., Head, J., Kivimäki, M., & Vahtera, J. (2014). Age-related trajectories of physical functioning in work and retirement: The role of sociodemographic factors, lifestyle and disease. *Journal of Epidemiology and Community Health*, *68*(6), 503–509. https://doi.org/10.1136/jech-2013-203555

Stock, J. H., & Yogo, M. (2002). *Testing for Weak Instruments in Linear IV Regression* (Working Paper 284; Technical Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/t0284

Syse, A., Veenstra, M., Furunes, T., Mykletun, R. J., & Solem, P. E. (2017). Changes in Health and Health Behavior Associated With Retirement. *Journal of Aging and Health*, *29*(1), 99–127. https://doi.org/10.1177/0898264315624906

Szabó, Á., Allen, J., Stephens, C., & Alpass, F. (2019). Is retirement associated with physical health benefits? A longitudinal investigation with older New Zealanders. *Age and Ageing*, *48*(2), 267–272. https://doi.org/10.1093/ageing/afy176

the United States Social Security Administration. (2020, March). *Social Security Programs Throughout the World*. https://www.ssa.gov/policy/docs/progdesc/ssptw/index.html

the World Bank. (2022, May 25). *World Development Indicators*. The World Bank. https://databank.worldbank.org/source/world-development-indicators/

Tierney, M. C., Moineddin, R., & McDowell, I. (2010). Prediction of all-cause dementia using neuropsychological tests within 10 and 5 years of diagnosis in a community-based sample. *Journal of Alzheimer's Disease: JAD*, *22*(4), 1231–1240. https://doi.org/10.3233/JAD-2010-100516

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Twomey, L. T., & Taylor, J. R. (1984). Old age and physical capacity: Use it or lose it. *The Australian Journal of Physiotherapy*, *30*(4), 115–120. https://doi.org/10.1016/S0004-9514(14)60683-8

United Kingdom Government. (2014, May 15). *State Pension age timetable*. GOV.UK. https://www.gov.uk/government/publications/state-pension-age-timetable

van Zon, S. K. R., Bültmann, U., Reijneveld, S. A., & de Leon, C. F. M. (2016). Functional health decline before and after retirement: A longitudinal analysis of the Health and Retirement Study. *Social Science & Medicine (1982)*, *170*, 26–34. https://doi.org/10.1016/j.socscimed.2016.10.002

Vélez-Coto, M., Andel, R., Pérez-García, M., & Caracuel, A. (2021). Complexity of work with people: Associations with cognitive functioning and change after retirement. *Psychology and Aging*, *36*, 143–157. https://doi.org/10.1037/pag0000584

Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, *113*(523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839

Welsh, K. A., Butters, N., Mohs, R. C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology*, *44*(4), 609–614. https://doi.org/10.1212/wnl.44.4.609

Wimo, A., Guerchet, M., Ali, G.-C., Wu, Y.-T., Prina, A. M., Winblad, B., Jönsson, L., Liu, Z., & Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *13*(1), 1–7. https://doi.org/10.1016/j.jalz.2016.07.150

Wittenberg, R., Hu, B., Barraza-Araiza, L., & Rehill, A. (2019). *Projections of older people with dementia and costs of dementia care in the United Kingdom, 2019–2040* (CPEC Working Paper 5). Care Policy and Evaluation Centre, London School of Economics and Political Science. https://www.lse.ac.uk/cpec/assets/documents/cpec-working-paper-5.pdf

World Health Organization. (2015). *World report on ageing and health*. World Health Organization. https://apps.who.int/iris/handle/10665/186463

World Health Organization. (2017). *Global action plan on the public health response to dementia 2017—2025*. http://www.who.int/mental_health/neurology/dementia/action_plan_2017_2025/en/

World Health Organization. (2020). *Decade of Healthy Ageing 2020-2030*. World Health Organization. https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action

World Health Organization. (2021, June 9). *Obesity and overweight*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Xue, B., Head, J., & McMunn, A. (2017). The Associations Between Retirement and Cardiovascular Disease Risk Factors in China: A 20-Year Prospective Study. *American Journal of Epidemiology*, *185*(8), 688–696. https://doi.org/10.1093/aje/kww166

Xue, B., Head, J., & McMunn, A. (2020). The Impact of Retirement on Cardiovascular Disease and Its Risk Factors: A Systematic Review of Longitudinal Studies. *The Gerontologist*, *60*(5), e367–e377. https://doi.org/10.1093/geront/gnz062

Zamarro, G., & Lee, J. (2012). *Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Employment and Retirement Measures*. RAND Corporation. https://www.rand.org/pubs/working_papers/WR861z4.html

Zhu, R. (2016). Retirement and its consequences for women's health in Australia. *Social Science & Medicine (1982)*, *163*, 117–125. https://doi.org/10.1016/j.socscimed.2016.04.003

Zotcheva, E., Bratsberg, B., Strand, B. H., Jugessur, A., Engdahl, B. L., Bowen, C., Selbæk, G., Kohler, H.-P., Harris, J. R., Weiss, J., Tom, S. E., Krokstad, S., Mekonnen, T., Edwin, T. H., Stern, Y., Håberg, A. K., & Skirbekk, V. (2023). Trajectories of occupational physical activity and risk of later-life mild cognitive impairment and dementia: The HUNT4 70+ study. *The Lancet Regional Health - Europe*, 100721. https://doi.org/10.1016/j.lanepe.2023.100721

Appendixes

A. Measurement of Diagnosed Health Conditions[36]

There was variation in the measurement of diagnosed health conditions across surveys. In the harmonized datasets, the variable of a general diagnosis of heart disease was constructed to maximize the number of compatible surveys rather than a specific condition such as myocardial infarction, angina, and congestive heart failure.

A.1 SHARE

SHARE asks participants to look at a card that says, "Please look at card 8. Has a doctor ever told you that you had/do you currently have any of the conditions listed on the card? With this, we mean that a doctor has told you that you have this condition and that you are either currently being treated for or bothered by this condition." The conditions included 1) a heart attack including myocardial infarction, coronary thrombosis, or any other heart problem, including congestive heart failure, 2) stroke or cerebral vascular disease, 3) high blood pressure or hypertension, and 4) diabetes or high blood sugar.

A.2 ELSA

ELSA shows the participants a card with a list of chronic medical conditions and asks whether a doctor has told them that they have a specific condition. The conditions included 1) angina, 2) heart attack (including myocardial infarction or coronary thrombosis), 3) congestive heart failure, 4) heart murmur, 5) abnormal heart rhythm, 6) stroke, 7) high blood pressure or hypertension, and 8) diabetes or high blood sugar.

A.3 CRELES

CRELES asks the participants whether a physician has told them that they have chronic medical conditions. The conditions included 1) heart attack, 2) other heart diseases, 3) stroke, 4) high blood pressure or hypertension, and 5) diabetes or high blood sugar levels.

A.4 MHAS

---

[36] Hu P, Lee J. Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Chronic Medical Conditions. RAND Corporation, 2012 https://www.rand.org/pubs/working_papers/WR861z1.html (accessed 26 April 2022).

MHAS asks the participants whether a doctor or medical personnel has ever told them that they have chronic medical conditions. The conditions included 1) heart attack, 2) stroke, 3) hypertension or high blood pressure, and 4) diabetes.

A.5 HRS

HRS asks the participants whether doctors have ever told them that they have chronic medical conditions. Medical doctors may include specialists such as dermatologists, psychiatrists, ophthalmologists, osteopaths, cardiologists, family doctors, internists, and physician assistants, but do not include chiropractors, dentists, or nurses/nurse practitioners. The conditions included 1) heart problems, including heart attack, coronary heart disease, angina, congestive heart failure, or other, 2) stroke, 3) hypertension, and 4) diabetes.

A.6 CHARLS

CHARLS asks the participants whether the chronic medical conditions have been diagnosed by health professionals or are known by the respondent themselves. It accepts diagnoses by physicians and other healthcare providers, including nurses, paramedics, and doctors of traditional Chinese medicine. The conditions included 1) heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems, 2) stroke, including transient ischaemic attack or TIA, 3) hypertension, and 4) diabetes or high blood sugar.

A.7 JSTAR

JSTAR asks the participants, “Please tell me if you have been diagnosed by a doctor or advised to seek care in connection with any of them.” Chronic medical conditions included 1) heart disease, 2) stroke, 3) high blood pressure, and 4) diabetes.

A.8 KLoSA

KLoSA asks the participants whether there has been a physician diagnosis of certain chronic medical conditions. The conditions included 1) heart disease, 2) stroke, 3) high blood pressure, and 4) diabetes.

B. Measurement of Risk Factors for Cardiovascular Disease

B.1 Obesity (BMI)

Obesity was defined as a BMI of 30 kg/m2 or higher.1F[37] BMI was calculated from self-reported height and weight, except for ELSA and CHARLS. In ELSA, a nurse measured participants' height and weight in waves 2, 4, and 6, while the most recent measurement of height and self-reported weight was used to calculate BMI in wave 8 (BMI is not provided in waves 1, 3, 5, 7, and 9). In CHARLS, height and weight were objectively measured in waves 1–3 but not in wave 4.

B.2 Physical inactivity

We considered those who engaged in vigorous or moderate physical activity less than once per week to be physically inactive. We constructed a binary variable from two original variables of the frequency of vigorous and moderate physical activities. However, for South Korea, we used only the variable of vigorous physical activity because the KLoSA did not ask about the frequency of moderate physical activity. In wave 7 of SHARE, only those who participated in wave 3 were asked about the frequency of physical activity; thus, all observations from Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania, and Slovakia were excluded from the analysis because individuals had only one observation. In waves 1–3 of CHARLS, only half of the participants were asked questions. We also excluded some observations due to the incompatibility of questions; MHAS, waves 1–2 of CRELES, and waves 1–6 of HRS asked whether the participant engaged in vigorous physical activity three times or more per week; and JSTAR asked for minutes of exercise on weekdays and weekends.

B.3 Smoking status

Smoking status indicates whether the participant is currently smoking. In Wave 6 of SHARE, those who had been interviewed previously were not asked about their smoking status. In wave 7, only new participants and those who were in wave 3 and had previously reported smoking were asked about their current smoking status. Thus, all the observations from Bulgaria, Cyprus, Finland, Latvia, Lithuania, Malta, Romania, and Slovakia were excluded due to a single observation.

B.4 Binge drinking

[37] World Health Organization. Obesity and overweight. World Health Organ. 2021; published online June 9. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (accessed April 27, 2022).

We defined binge drinking as consuming five or more drinks per day for men and four or more for women.[38] Waves 1 and 6–8 of SHARE, wave 1 of ELSA, waves 1–2 of HRS, and all waves of CRELES and CHARLS did not provide the number of drinks per day. Thus, all the observations from Bulgaria, Croatia, Cyprus, Finland, Greece, Hungary, Latvia, Lithuania, Luxembourg, Malta, Portugal, Romania, and Slovakia were excluded from the analysis because individuals had only one observation. In some study sites of JSTAR, the question was not asked in wave 2.

C. Measurement of Labor Force Status[39]

C.1 SHARE

SHARE asks participants, "In general, how would you describe your current situation?" They then choose the best description of their current labor force status from a list of options: 1) retired, 2) employed or self-employed (including working for a family business), 3) unemployed and looking for work, 4) permanently sick or disabled, 5) homemaker, or 6) other (renter, living off own property, student, or doing voluntary work). The harmonized variable was constructed based on responses to this direct question.

C.2 ELSA

ELSA asks participants, "Which of these, would you say, best describes your situation?" They then choose the best description of their current labor force status from a list of options: 1) employed, 2) self-employed, 3) unemployed, 4) partly retired, 5) retired, 6) permanently sick or disabled, or 7) looking after home or family. The harmonized variable was constructed based on responses to this direct question.

C.3 CRELES

CRELES asks participants whether they have ever had a job for which they received payment in money or kind. If the respondent answered yes, they were asked what they did during most of the last week: worked, worked with a family business, did not work but had a job, looking

---

[38] Centers for Disease Control and Prevention. Binge Drinking [Internet]. Cent. Dis. Control Prev. 2022 [cited 2022 Jun 16]. Available from: https://www.cdc.gov/alcohol/fact-sheets/binge-drinking.htm
[39] Zamarro G, Lee J. Harmonization of Cross-National Studies of Aging to the Health and Retirement Study: Employment and Retirement Measures. RAND Corporation, 2012 https://www.rand.org/pubs/working_papers/WR861z4.html (accessed 28 April 2022).

for work, did household chores, or did not work. Participants who were not working were then asked if they had not worked: less than 2 years, more than 2 years, or had never worked.

If the participant answered that they worked, helped with a family business, or did not work last week but had a job, the harmonized variable was set to “working.” If the participant reported that they were looking for work, it was set to “unemployed.” If the participant reported that they were doing household chores, it was set to “doing household chores.” If the participant reported that they had worked in the past but were not currently working, it was set to “retired.” If the participant reported that they had never worked, it was set to “never worked.”

C.4 MHAS

MHAS asks participants whether they have worked or are currently working. In waves 2–4, it also asks the main reason why they were not currently working: dedicated to household chores, retired, old age, sick or temporarily disabled, unable to work for the rest of life, and does not have customers or cannot find work.

If the participant reported that they were currently working, the harmonized variable was set to “working.” If the participant is currently looking for work or does not work but “does not have customers or cannot find work,” it is set to “unemployed.” If the participant mentioned retirement, regardless of current work, it was set to “retired.” If the participant is “sick or temporarily disabled” or “unable to work for rest of life,” it is set to “disabled.” Otherwise, the variable is set to “not in the labor force.” The question asking the reason for not working is not included in wave 1; thus, the harmonized variable has an integrated category indicating “unemployed, retired, or disabled” in wave 1. If the present study, we treated those in the integrated category in wave 1 as retirees only if they were categorized as “retired” in wave 2; otherwise, they were not included in the analyses.

C.5 HRS

Participants in HRS provided information on their labor force status at several time points in an interview. First, HRS asks the participants to select all applicable options from a list that includes 1) working now, 2) unemployed and looking for work, 3) temporarily laid off, on sick or other leave, 4) disabled, 5) retired, 6) homemaker, or 7) other (specify). It also asks them whether they are currently working for payment, the usual number of hours per week if applicable, and whether they consider themselves partly retired, completely retired, or not retired.

If the participant reports working full-time (i.e., working 35+ hours per week or 36+ weeks per year), the harmonized variable is set to "working full-time." If the participant is working part-time and does not mention retirement, it is set to "working part-time." If the participant is working part-time and mentions retirement, it is set to "partly retired." If the participant is not working but is looking for a job, it is set to "unemployed." If the participant is not looking for a job and there is any mention of retirement, it is set to "retired." If retirement is not mentioned and disabled employment status is given, it is set to "disabled." Otherwise, the variable is set to "not in the labor force."

C.6 CHARLS

CHARLS asks participants whether they engaged in agricultural work for more than 10 days in the past year, worked for at least 1 hour last week if not engaged in agricultural work, were temporarily laid-off or on sick or other leave, worked for at least a few months, whether they were homemakers, completed retirement procedures, and currently retired (including early retirement or internal retirement).

If the participants report working for other farmers, the harmonized variable is set to "agricultural employed." If the participants reported working for their household, it is set to "agricultural self-employed." If the participants describe their non-agricultural job as employment, it is set to "non-agricultural employed." If the participants describe their non-agricultural job as self-employment, it is set to "non-agricultural self-employed." If the participants describe their non-agricultural job as an unpaid family business, it is set to "non-agricultural unpaid family business." If the participants report not currently working but had worked for at least 3 months and have searched for a job in the past month, it is set to "unemployed." If the participants declare to have completed retirement procedures or describe themselves as retired, it is set to "retired." If the participants reported never worked, it was set to "never worked."

C.7 JSTAR

JSTAR asks participants whether they are currently employed, looking for a job, or intend to look for work in the future. If they were neither a worker nor a job seeker, they were asked about their current status with the following response options: 1) retired, 2) keep house, 3) receive medical care, 4) other, 5) do not know, and 6) refused to answer.

If the participant reports working full-time or working 35+ hours per week and 36+ weeks per year, the harmonized variable is set to "working full-time." If the participant reports working part-time or less than 35 hours per week or 36 weeks per year, it is set to "working part-time." If the participant reports currently working as an owner of an independent business or having a side job at home, it is set to "self-employed."

If the participant reports not working but is looking for a job and there is no mention of retirement, it is set to "unemployed." If the participant reports looking for a part-time job and mentions retirement, it is set to "partly retired." If the participant is not working and not looking for work, and there is any mention of retirement, it is set to "retired." If retirement is not mentioned and disabled employment status is given, it is set to "disabled." If neither retirement nor disability is mentioned, but a homemaker situation is given, it is set to "not in the labor force."

C.8 KLoSA

KLoSA asks participants whether they are currently working or looking for a job. If they are neither a worker nor a job seeker, they are asked about their retirement status with the following response options: 1) worked before but currently retired, 2) worked before and intended to work in the future but currently not looking for a job, and 3) never had a job before.

If the participant is employed by another person or company for payment, the harmonized variable is set to "employed full-time" or "employed part-time," based on the working classification the participant gave for the job. If the participant report being self-employed, it is set to "self-employed." If the participant is employed and reports working without payment for family more than 18 hours per week, it is set to "help with family 18 hours or more per week." If a non-working participant is looking for work and reports being able to work if offered a job and then confirms that they have done something to find work in the last 4 weeks, it is set to "unemployed." If a non-working participant is not looking for work and reports being retired, it is set to "retired." If the participant reports being retired but later mentions working for payment or looking for paid work, it is set to "partly retired." If a non-working participant is looking for work but then reports not being able to accept work or looking for work due to poor health or a disability, it is set to "disabled." Otherwise, it is set to "not in labor force."

Table C.1 Summary of the Harmonized Variable of Labor Force Status

This study	**SHARE**	**ELSA**	**CRELES**	**MHAS**	**HRS**	**CHARLS** (urban residents only)	**JSTAR**	**KloSA**
Included as those "working"	1. employed or self employed	1. employed	1. working	1. working	1. working full-time	1. agricultural employed	1. working full-time	1. working full-time
		2. self-employed			2. working part-time	2. agricultural self-employed	2. working part-time	2. working part-time
						3. non-agricultural employed	8. self-employed	3. self-employed
						4. non-agricultural self-employed		4. help with family 18 hours or more per week
						5. non-agricultural unpaid family business		
Included as those being "retired"	5. retired	4. partly retired	4. retired	3. retired	4. partly retired	7. retired	4. partly retired	6. partly retired
		5. retired			5. retired		5. retired	7. retired
Excluded from analyses	3. unemployed	3. unemployed	2. unemployed	2. unemployed	3. unemployed	6. unemployed	3. unemployed	5. unemployed
	6. permanently sick or disabled	6. disabled	3. doing household chores	4. disabled	6. disabled	8. never worked	6. disabled	8. disabled
	8. homemaker	7. looking after home or family	5. never worked	5. not in labor force	7. not in labor force		7. not in labor force	9. not in labor force

Retirement status was determined based on the harmonized variable of labor force status (RwLBRF).

D. Early and Official Retirement Age by Country

Table D.1 Early and Official Retirement Age by Country

Country	Year	Men		Women	
		ERA	ORA	ERA	ORA
Austria[a]	2018	NA	65	NA	60
Belgium[b]	2018	63	65	63	65
Bulgaria[c]	2018	63.08	64.08	60.17	61.17
Croatia[d]	2018	60	65	57	62
Cyprus	2018	63	65	63	65
Czech Republic[e]	2018	60	63.16	59.66	62.66
Denmark[f]	2018	NA	65	NA	65
England[g]	2018	NA	65	NA	65
Estonia[h]	2018	60.5	63.5	60.5	63.5
Finland	2018	63	65	63	65
France[i]	2018	62	67	62	67
Germany[j]	2018	63	65.58	63	65.58
Greece[k]	2018	62	67	62	67
Hungary[l]	2018	NA	63.5	NA	63.5
Israel[m]	2018	NA	67	NA	62
Italy[n]	2018	62	66.58	62	66.58
Latvia[o]	2018	61.25	63.25	61.25	63.25
Lithuania[p]	2018	58.67	63.67	57.33	62.33
Luxembourg	2018	57	65	57	65
Malta[q]	2018	61	62	61	62
Netherlands[r]	2018	NA	66	NA	66
Poland[s]	2018	NA	65	NA	60
Portugal[t]	2018	60	66.33	60	66.33
Romania[u]	2018	60	65	55.92	60.92
Slovakia[v]	2018	60.42	62.42	60.42	62.42
Slovenia[w]	2018	60	65	59.67	64
Spain[x]	2018	61.5	65.5	61.5	65.5
Sweden[y]	2018	61	65	61	65
Switzerland	2018	63	65	62	64
Costa Rica	2013	62	65	60	65
Mexico	2018	60	65	60	65
United States[z]	2018	62	66	62	66
China[aa]	2016	NA	60	NA	50
Japan	2012	60	65	60	65
South Korea[ab]	2018	57	61	57	61

Source: The United States Social Security Administration “Social Security Programs Throughout the World”; Organisation for Economic Co-operation and Development “Pensions at a Glance”; websites of the authorities of each country.

Note: ERA, ORA, and NA denote early retirement age, official retirement age, and not applicable, respectively.

[a] ERA was 61.5 for men and 56.5 for women in 2004 and gradually increased to be phased out in 2017.

[b] ERA gradually increased from age 60 to 63 from 2013 to 2018.

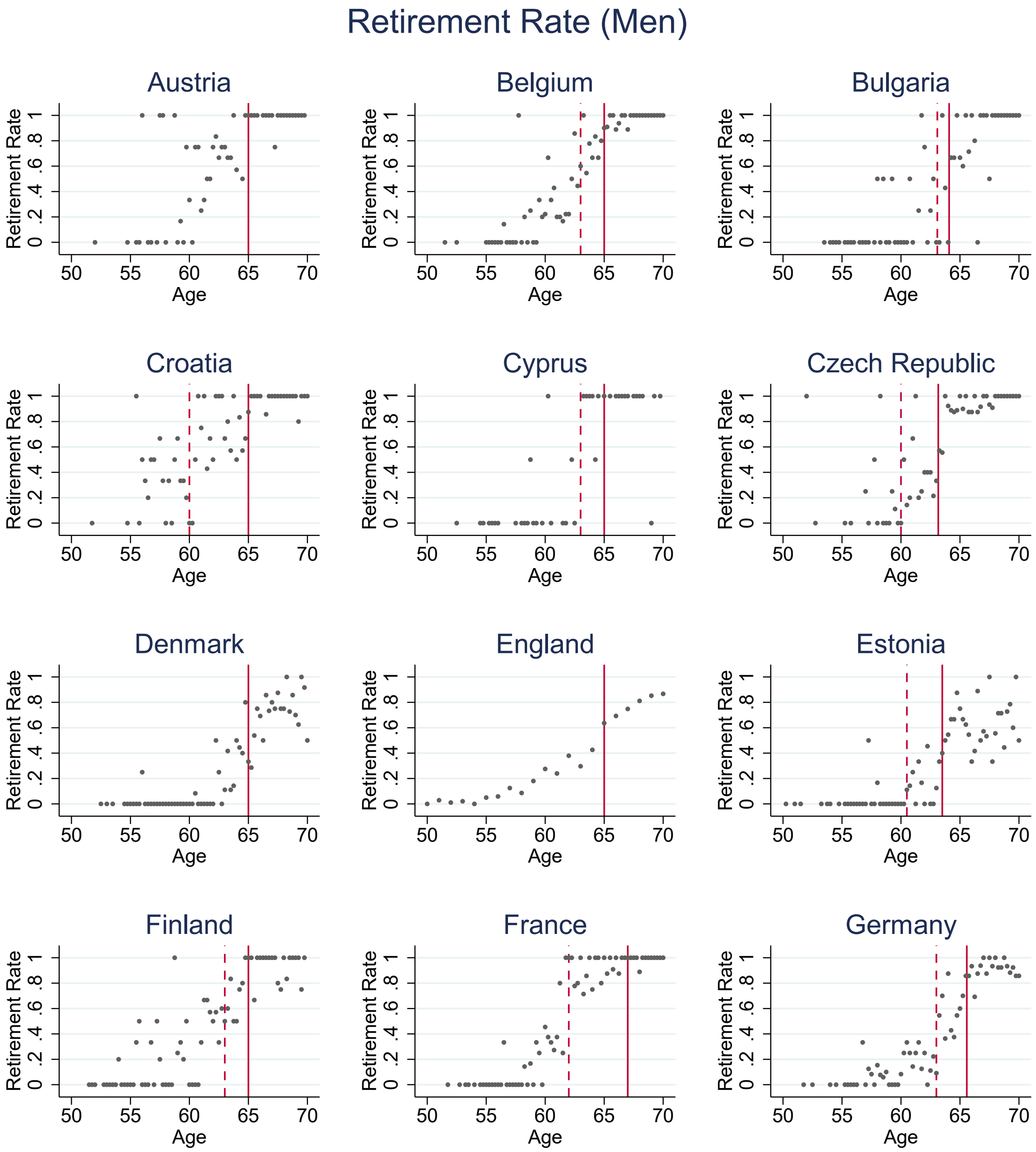
[c] ORA is gradually increasing from age 63 to 65 by 2029 for men and from age 60 to 65 by 2037 for women. Early retirement is possible up to one year prior to the ORA.

[d] ERA and ORA for women are gradually increasing from age 55 to 60 and 60 to 65 by 2030, respectively.
[e] ORA is gradually increasing from age 60 to 65 for men and 57 to 65 for women without children by 2030.
[f] ORA was 67 for those who reached age 60 before 1 July 1999. ORA is gradually increasing to age 67 from 2019 to 2022.
[g] ORA for women gradually increased from age 60 to 65 from 2010 to 2018.
[h] ORA is gradually increasing from age 63 to 65 from 2017 to 2026. Early retirement is possible up to three years prior to the ORA.
[i] ERA is increasing from age 60 to 62, and ORA from 65 to 67, depending on the year of birth.
[j] ERA and ORA are gradually increasing from age 63 to 65 and 65 to 67 by 2029, respectively.
[k] ERA increased from age 60 to 62 for men and 55 to 62 for women, and ORA from 65 to 67 for men and 60 to 67 for women in 2013.
[l] ORA is gradually increasing from age 62 to 65 by 2022.
[m] ORA is increasing from age 65 to 67 for men and 60 to 62 for women, depending on the year of birth.
[n] ERA gradually increased from age 57 to 62, and ORA from 65 to 67 for men and 60 to 67 for women in 2019.
[o] ORA is gradually increasing from age 62 to 65 from 2013 to 2025. Early retirement is possible up to two years prior to the ORA.
[p] ORA is gradually increasing to age 65 by 2026. Early retirement is possible up to five years prior to the ORA.
[q] ORA is gradually increasing to age 65, depending on the year of birth.
[r] ORA is gradually increasing to age 67 by 2024, depending on the year of birth.
[s] ORA increased from age 65 to 65.58 for men and 60 to 60.58 for women from 2012 to 2015 but returned to age 65 and 60 in 2017.
[t] ERA increased from age 55 to 60 in 2015, and ORA is gradually increasing to age 66.5 by 2021.
[u] ORA for women is gradually increasing to age 63 by 2030. Early retirement is possible up to five years prior to the ORA.
[v] ORA is gradually increasing from age 62 based on increases in life expectancy from 2016. Early retirement is possible up to two years prior to the ORA.
[w] ERA (with 40 years of contribution) gradually increased from age 58 to 60 in 2018 for men and 2019 for women. ORA (with 20 years of contribution) gradually increased from age 63 to 65 from 2012 to 2016 for men and 61 to 65 from 2012 to 2020 for women.
[x] ORA is gradually increasing from age 65 to 67 from 2012 to 2027. Early retirement is possible up to four years prior to the ORA in the case of involuntary unemployment.
[y] The earning-related national pension and guarantee pension benefits are available from ages 61 and 65, respectively.
[z] ORA gradually increased from age 65 to 66, depending on the year of birth.
[aa] ORA of 50 is for non-professional salaried women.
[ab] ERA is gradually increasing from age 55 to 60 from 2012 to 2029, and ORA from 60 to 65 from 2012 to 2034.

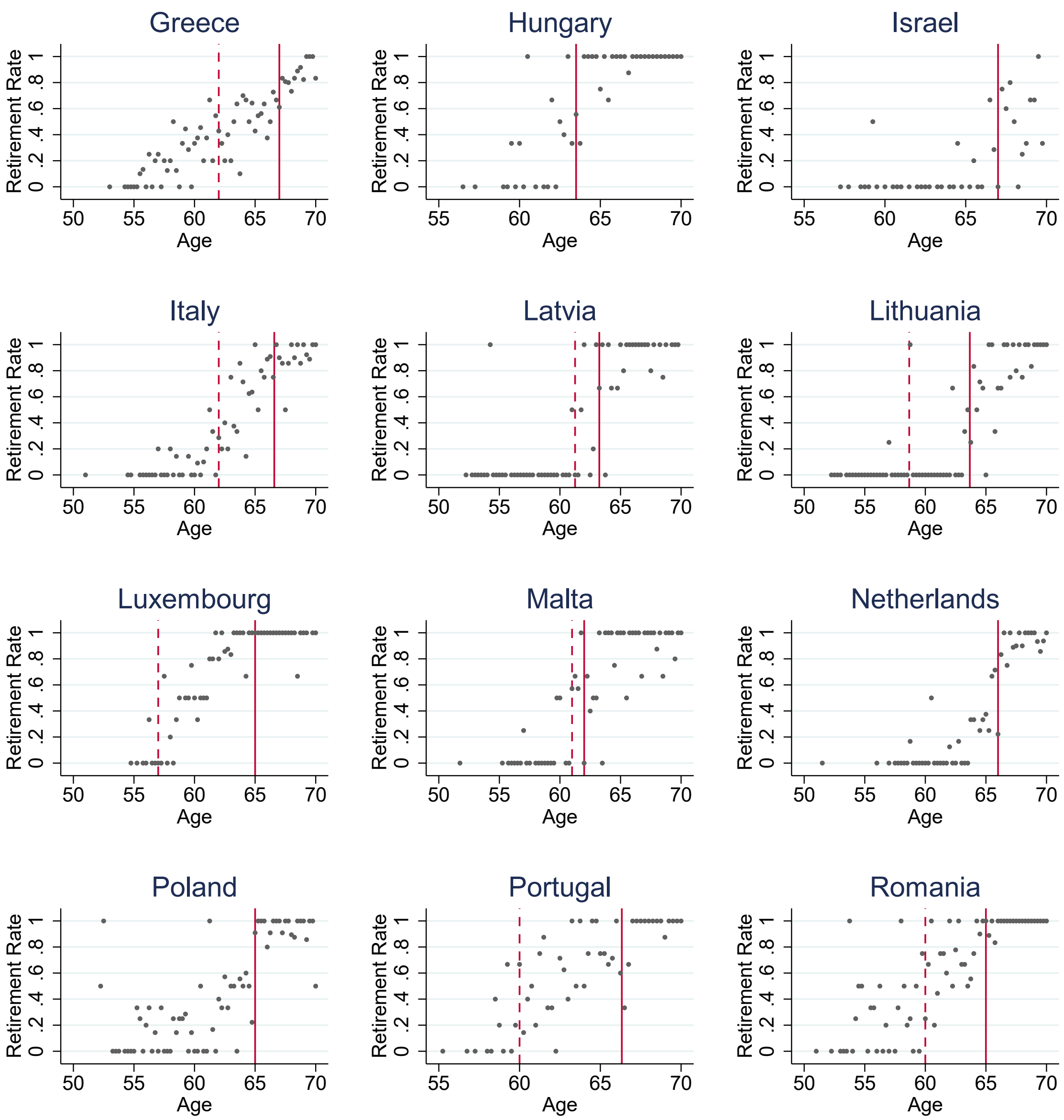
E. Retirement Rate by Country

The latest surveys were utilized to illustrate the figures. Each individual dot represents the average retirement rate for each 3-month intervals, as monthly age data was unavailable in England and Japan. The retirement rate is calculated by dividing the number of retirees by the sum of retirees and non-retired individuals who are not working due to reasons other than retirement (such as being unemployed, disabled, or homemaker], with the exclusion of the latter. The dashed red line denotes the ERA, while the solid red line represents the ORA that corresponds to the survey year.

Figure E.1 Men’s Retirement Rate by Country



Retirement Rate (Men, cont.)



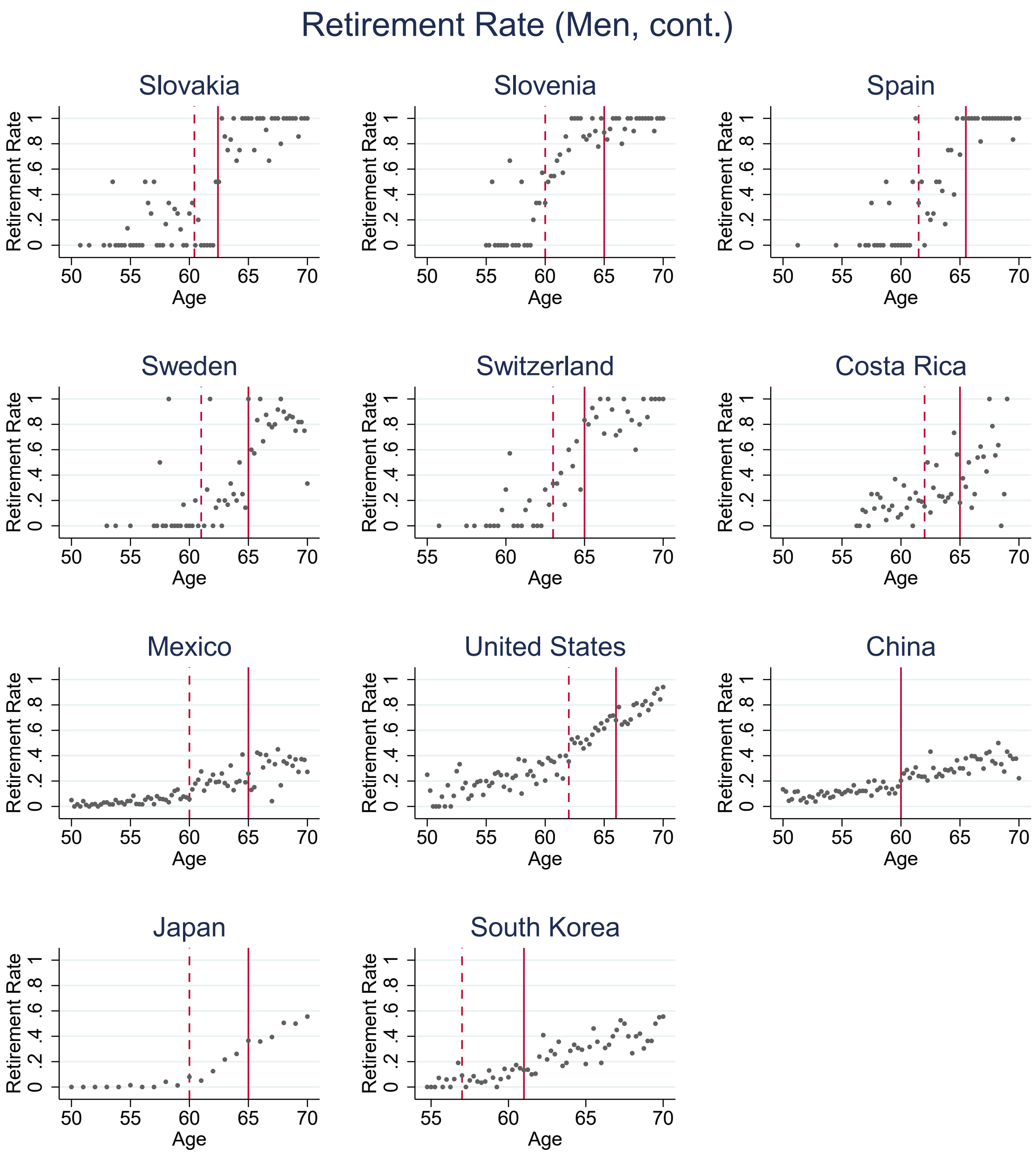
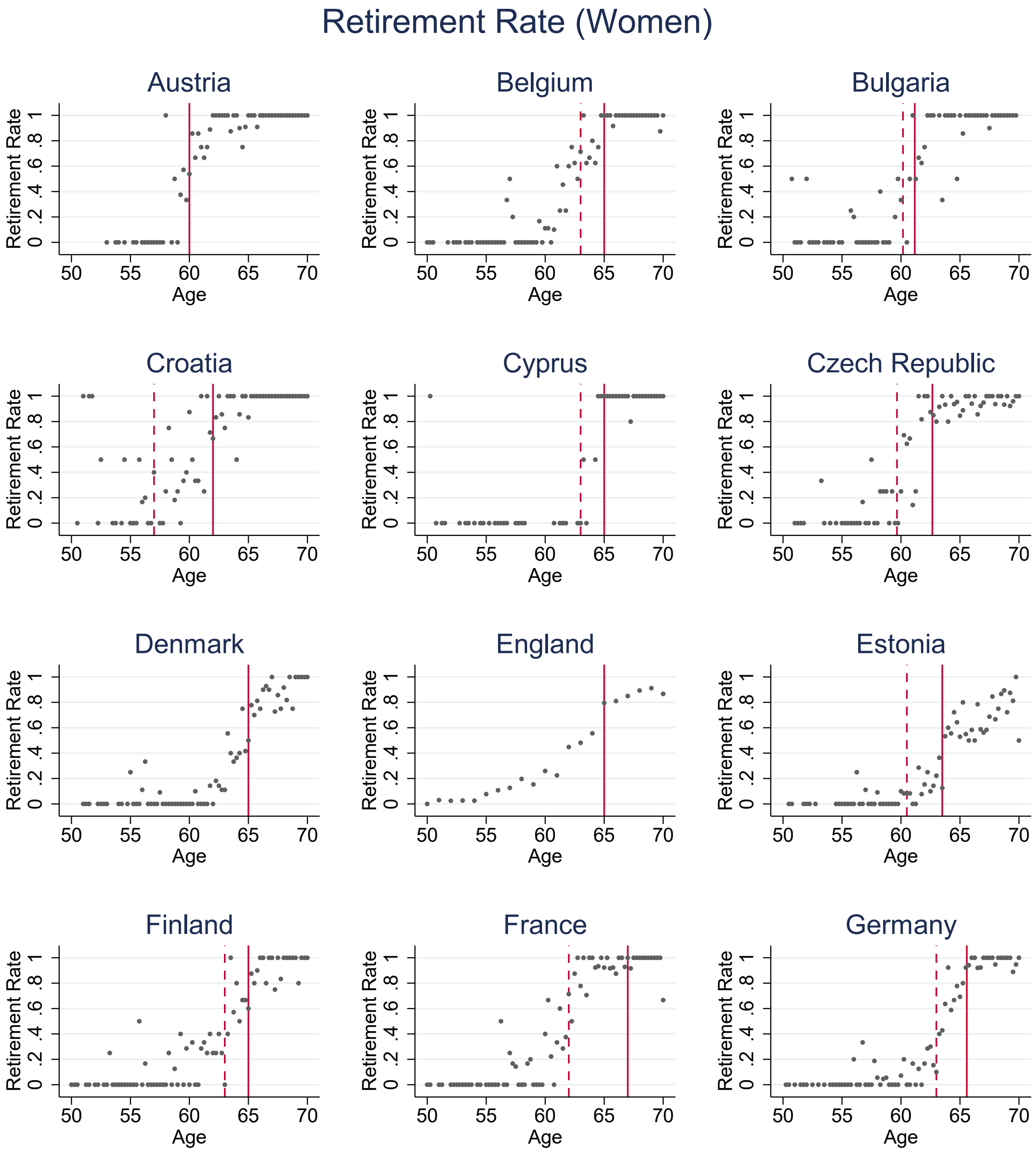
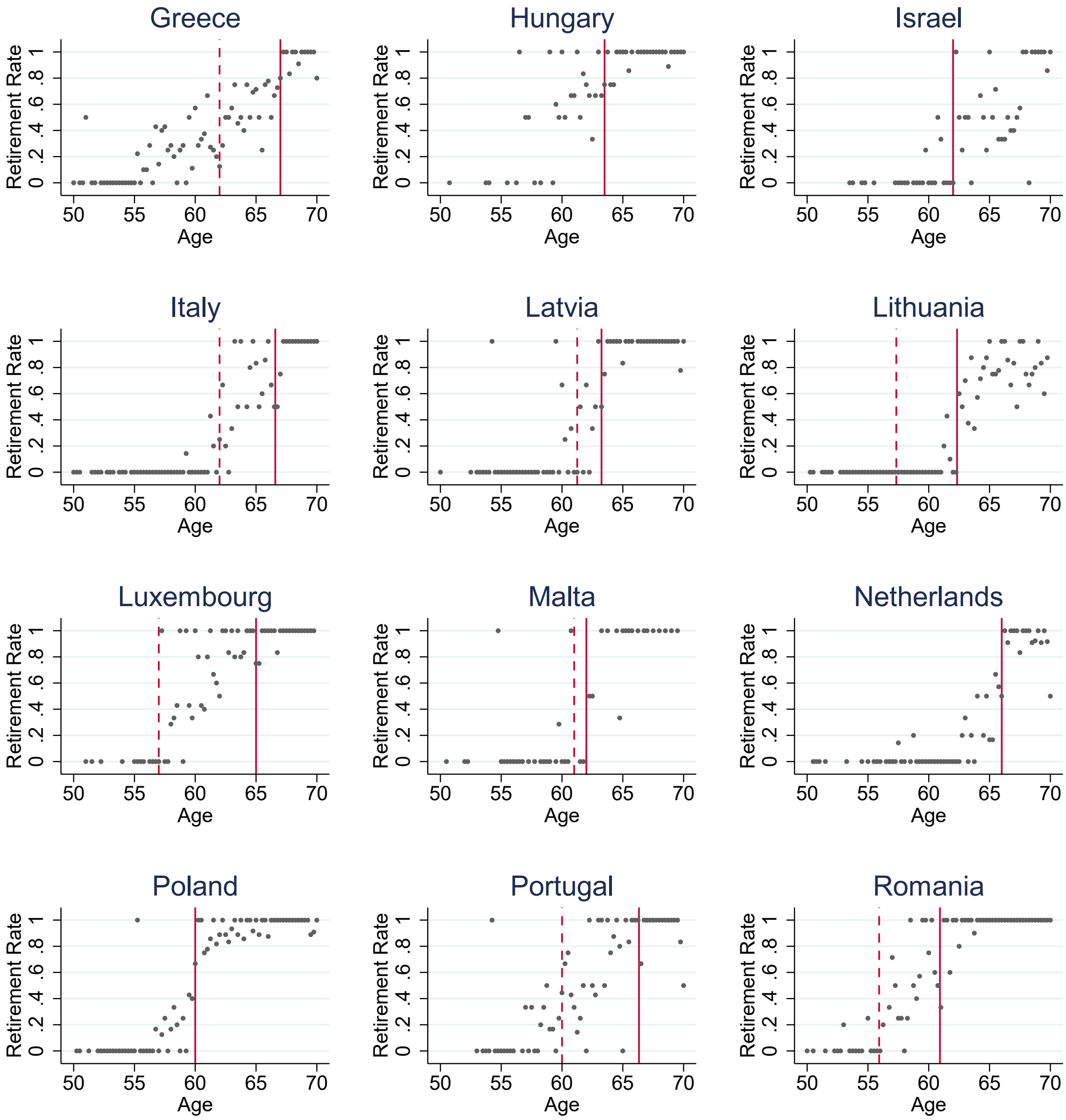

Retirement Rate (Men, cont.)
Slovakia
Slovenia
Spain
Sweden
Switzerland
Costa Rica
Mexico
United States
China
Japan
South Korea
Retirement Rate
Age

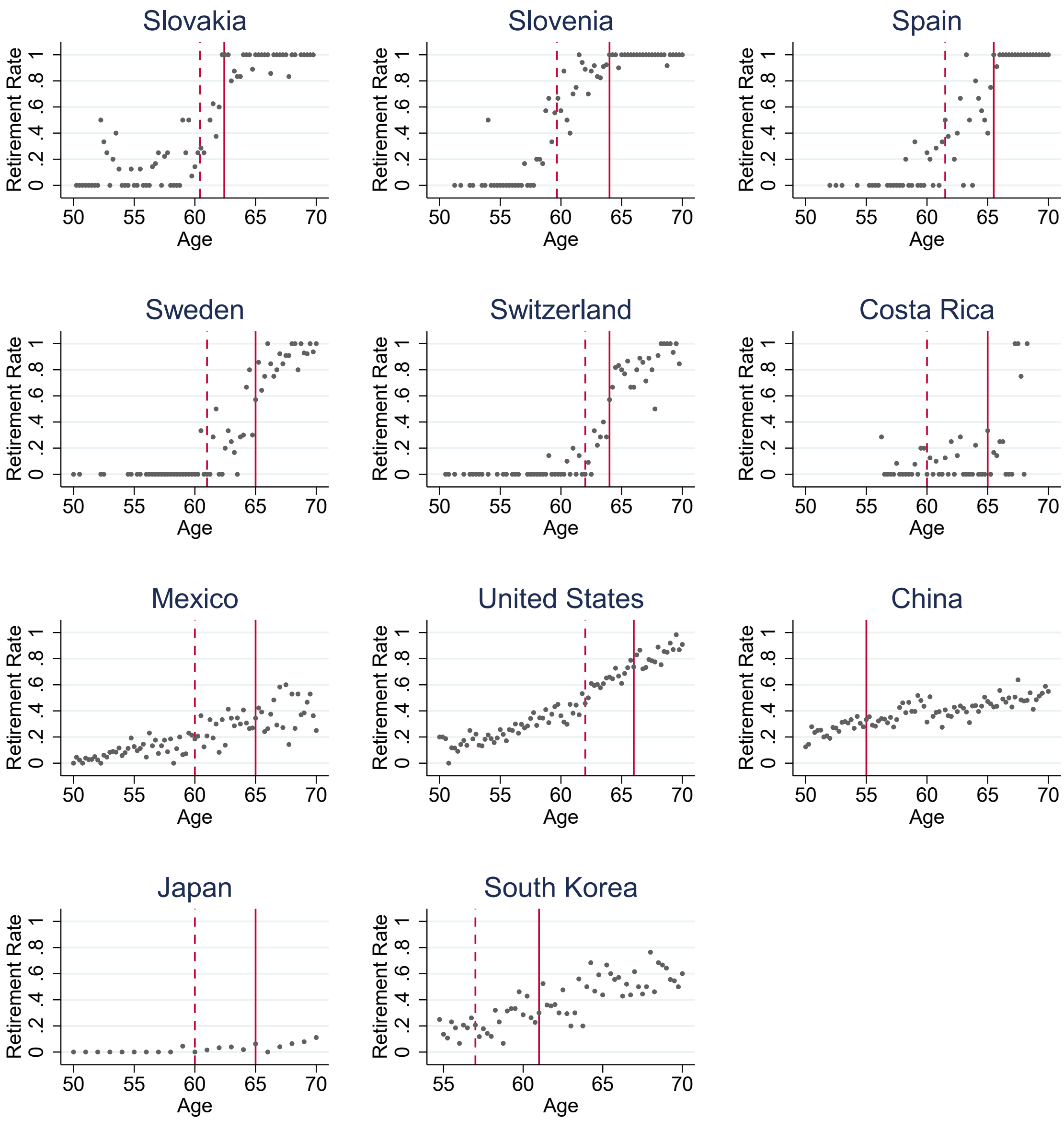
Figure E.2 Women's Retirement Rate by Country



Retirement Rate (Women, cont.)



Retirement Rate (Women, cont.)



F. Supplementary Materials for Chapter 1

Table F.1 Characteristics Comparison Between Followed Up and Lost to Follow-Up Participants

Characteristics in the previous interview	Lost to follow-up		Followed up		Standardised difference [a]
	Mean (N)	SD (%)	Mean (N)	SD (%)	
Retired	0.49	0.50	0.44	0.50	0.092
Age	61.28	5.41	60.44	5.51	0.154
Men	0.52	0.50	0.49	0.50	0.061
Education					0.098
Low	(13 700)	(31.7)	(79 318)	(27.4)	
Middle	(19 175)	(44.3)	(139 831)	(48.2)	
High	(10 407)	(24.0)	(70 693)	(24.4)	
Married	0.78	0.41	0.78	0.41	0.001
Heart disease	0.14	0.35	0.12	0.32	0.077
Stroke	0.05	0.21	0.03	0.18	0.063
Hypertension	0.44	0.50	0.40	0.49	0.085
Diabetes	0.15	0.36	0.13	0.34	0.055
Obesity	0.23	0.42	0.25	0.43	-0.051
Physical inactivity	0.18	0.38	0.19	0.39	-0.026
Smoking	0.23	0.42	0.20	0.40	0.072
Binge drinking	0.08	0.28	0.09	0.28	-0.003

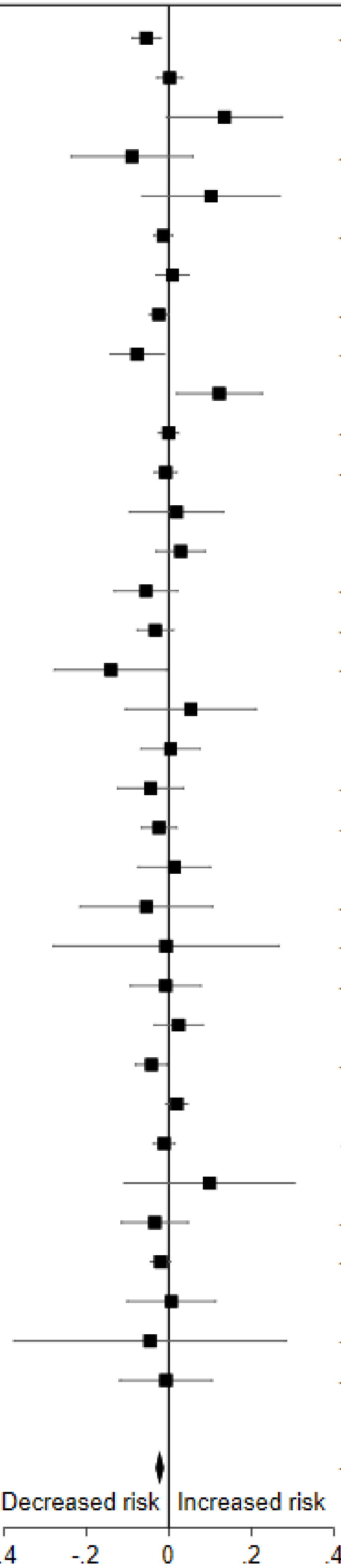
In general, standardised differences less than 0.1 indicate a balance between the two groups. [40,41]

[40] Cohen J. The t test for means. In: Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, N.J: Routledge; 1988. p. 19–74.
[41] Normand SLT, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, et al. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. J Clin Epidemiol. 2001 Apr 1;54(4):387–98.

Figure F.1 Association Between Retirement and Heart Disease by Country

Heart Disease

Country	Events	Obs.	Coef. (95% CI)	F-statistic
Austria	1001	9387	-0.054 (-0.088, -0.020)	123.671
Belgium	1299	13498	0.002 (-0.028, 0.033)	242.380
Bulgaria	103	748	0.135 (-0.005, 0.274)	11.939
Croatia	350	2627	-0.089 (-0.235, 0.058)	16.406
Cyprus	32	248	0.102 (-0.064, 0.269)	15.848
Czech Republic	1792	12865	-0.013 (-0.035, 0.009)	411.622
Denmark	926	10682	0.009 (-0.030, 0.049)	310.829
England	5898	44740	-0.024 (-0.046, -0.002)	765.038
Estonia	2125	11740	-0.076 (-0.141, -0.011)	54.595
Finland	92	1094	0.123 (0.019, 0.226)	21.678
France	1311	12052	-0.000 (-0.023, 0.023)	446.931
Germany	1110	10874	-0.007 (-0.034, 0.020)	308.897
Greece	526	5992	0.019 (-0.094, 0.133)	10.015
Hungary	267	1830	0.029 (-0.030, 0.089)	79.152
Israel	491	4695	-0.055 (-0.132, 0.022)	54.675
Italy	879	9658	-0.032 (-0.075, 0.011)	97.003
Latvia	72	598	-0.140 (-0.276, -0.005)	7.520
Lithuania	148	1052	0.054 (-0.104, 0.211)	15.616
Luxembourg	231	2355	0.005 (-0.067, 0.076)	55.227
Malta	26	476	-0.044 (-0.124, 0.036)	10.557
Netherlands	458	4934	-0.023 (-0.066, 0.019)	282.227
Poland	734	4558	0.014 (-0.073, 0.101)	69.065
Portugal	185	1756	-0.054 (-0.214, 0.106)	7.658
Romania	115	1112	-0.006 (-0.279, 0.267)	3.599
Slovakia	44	1320	-0.007 (-0.092, 0.078)	15.340
Slovenia	977	7791	0.024 (-0.034, 0.083)	81.499
Spain	794	8158	-0.042 (-0.080, -0.004)	200.253
Sweden	1016	10236	0.020 (-0.006, 0.046)	259.690
Switzerland	486	7890	-0.011 (-0.036, 0.014)	178.526
Costa Rica	80	2615	0.099 (-0.107, 0.305)	1.121
Mexico	813	21607	-0.033 (-0.114, 0.047)	15.192
United States	20247	133178	-0.019 (-0.043, 0.004)	397.573
China	1782	7680	0.006 (-0.100, 0.113)	32.666
Japan	354	4379	-0.045 (-0.374, 0.284)	1.841
South Korea	1082	21527	-0.006 (-0.118, 0.105)	7.250
All Countries (I^2 = 23.3%, 0.0–49.5)			-0.022 (-0.031, -0.012)	

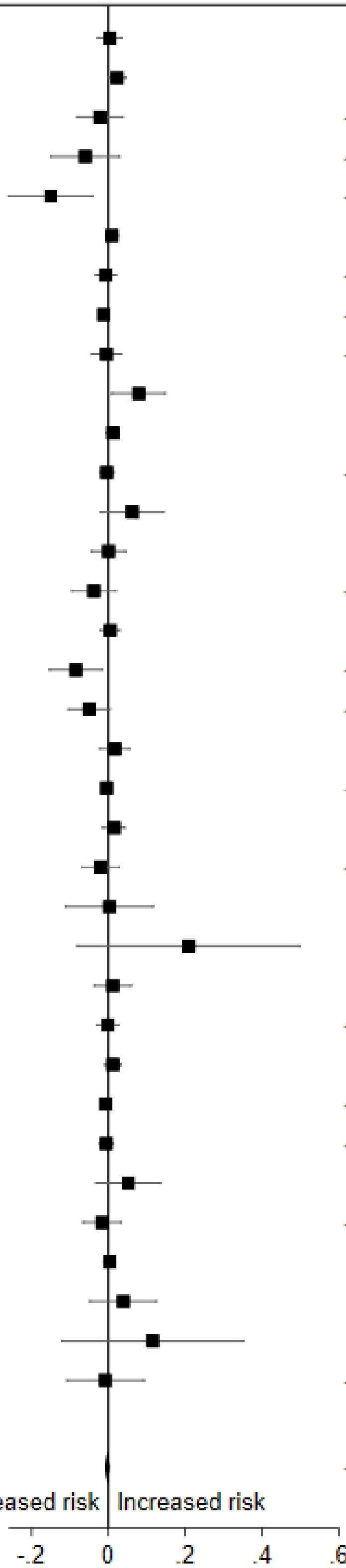


Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.2 Association Between Retirement and Stroke by Country

Stroke

Country	Events	Obs.	Coef. (95% CI)	F-statistic
Austria	1000	9387	0.004 (-0.028, 0.035)	123.273
Belgium	1297	13496	0.023 (0.001, 0.046)	242.368
Bulgaria	103	748	-0.021 (-0.081, 0.038)	11.939
Croatia	350	2627	-0.060 (-0.149, 0.028)	16.406
Cyprus	32	248	-0.150 (-0.260, -0.039)	15.848
Czech Republic	1789	12864	0.009 (-0.006, 0.024)	412.073
Denmark	925	10681	-0.006 (-0.032, 0.020)	310.833
England	5898	44740	-0.012 (-0.023, -0.001)	765.038
Estonia	2124	11739	-0.004 (-0.043, 0.034)	54.585
Finland	92	1094	0.079 (0.010, 0.148)	21.678
France	1310	12051	0.012 (-0.004, 0.028)	446.918
Germany	1108	10872	-0.004 (-0.021, 0.014)	309.174
Greece	525	5991	0.062 (-0.019, 0.144)	10.018
Hungary	267	1832	0.001 (-0.044, 0.045)	79.220
Israel	491	4695	-0.037 (-0.094, 0.020)	54.675
Italy	877	9656	0.004 (-0.020, 0.029)	97.013
Latvia	72	598	-0.084 (-0.153, -0.016)	7.520
Lithuania	148	1052	-0.049 (-0.104, 0.005)	15.616
Luxembourg	230	2354	0.016 (-0.022, 0.055)	55.235
Malta	26	476	-0.004 (-0.010, 0.002)	10.557
Netherlands	458	4934	0.015 (-0.014, 0.043)	282.227
Poland	734	4558	-0.020 (-0.068, 0.028)	69.065
Portugal	185	1756	0.003 (-0.111, 0.116)	7.658
Romania	115	1112	0.209 (-0.081, 0.499)	3.599
Slovakia	44	1320	0.011 (-0.036, 0.059)	15.340
Slovenia	976	7790	-0.001 (-0.030, 0.027)	81.469
Spain	794	8158	0.012 (-0.009, 0.033)	200.253
Sweden	1016	10236	-0.007 (-0.018, 0.005)	259.690
Switzerland	486	7890	-0.005 (-0.023, 0.013)	178.526
Costa Rica	80	2621	0.052 (-0.031, 0.135)	1.156
Mexico	813	21610	-0.016 (-0.065, 0.033)	15.217
United States	20247	133178	0.003 (-0.011, 0.018)	397.573
China	1758	7675	0.039 (-0.048, 0.125)	33.458
Japan	354	4375	0.116 (-0.120, 0.352)	1.866
South Korea	1082	21523	-0.008 (-0.108, 0.092)	7.240
All Countries (I^2 = 16.2%, 0.0–44.5)			-0.002 (-0.008, 0.004)	

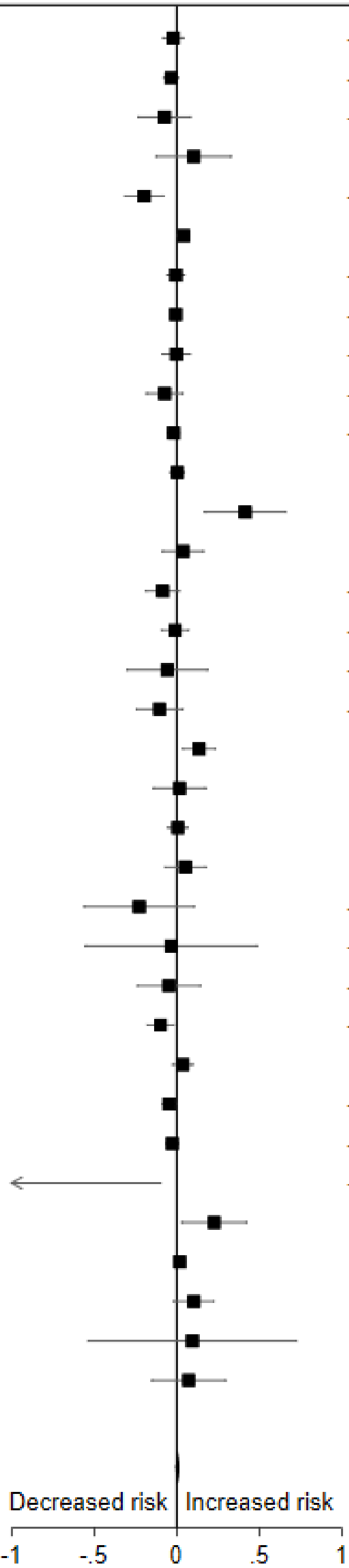


Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.3 Association Between Retirement and Hypertension by Country

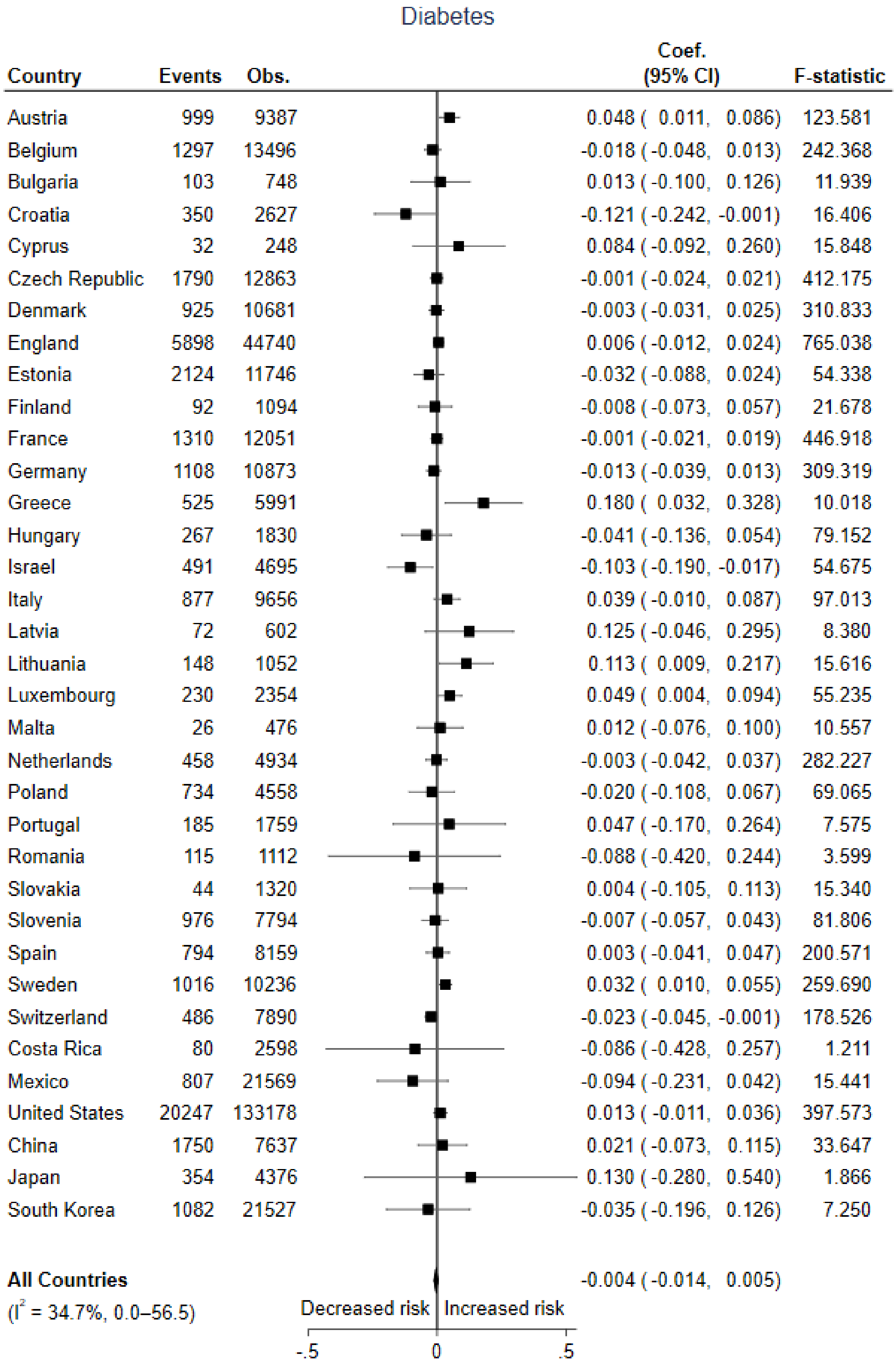
Hypertension

Country	Events	Obs.	Coef. (95% CI)	F-statistic
Austria	1000	9390	-0.020 (-0.081, 0.042)	123.582
Belgium	1299	13498	-0.033 (-0.078, 0.012)	242.380
Bulgaria	103	752	-0.074 (-0.232, 0.084)	12.548
Croatia	350	2627	0.103 (-0.123, 0.330)	16.406
Cyprus	32	248	-0.196 (-0.313, -0.080)	15.848
Czech Republic	1791	12867	0.043 (0.009, 0.078)	411.630
Denmark	926	10683	-0.004 (-0.056, 0.047)	311.423
England	5898	44740	-0.006 (-0.031, 0.020)	765.038
Estonia	2124	11748	-0.002 (-0.085, 0.082)	54.046
Finland	92	1094	-0.073 (-0.182, 0.036)	21.678
France	1310	12052	-0.018 (-0.050, 0.014)	447.330
Germany	1109	10873	0.004 (-0.038, 0.046)	309.162
Greece	526	5992	0.415 (0.171, 0.660)	10.015
Hungary	267	1833	0.039 (-0.085, 0.162)	79.282
Israel	491	4697	-0.084 (-0.186, 0.017)	55.140
Italy	877	9658	-0.009 (-0.089, 0.071)	97.022
Latvia	72	600	-0.055 (-0.299, 0.188)	8.388
Lithuania	148	1052	-0.103 (-0.242, 0.036)	15.616
Luxembourg	231	2357	0.135 (0.037, 0.234)	55.178
Malta	26	476	0.018 (-0.140, 0.176)	10.557
Netherlands	458	4934	0.007 (-0.050, 0.065)	282.227
Poland	734	4558	0.054 (-0.069, 0.176)	69.065
Portugal	185	1762	-0.226 (-0.559, 0.107)	7.641
Romania	115	1112	-0.033 (-0.551, 0.486)	3.599
Slovakia	44	1320	-0.046 (-0.236, 0.143)	15.340
Slovenia	976	7794	-0.098 (-0.177, -0.018)	81.801
Spain	794	8162	0.036 (-0.024, 0.096)	200.811
Sweden	1016	10236	-0.045 (-0.082, -0.008)	259.690
Switzerland	486	7890	-0.027 (-0.063, 0.010)	178.526
Costa Rica	80	2606	-1.071 (-2.045, -0.097)	1.158
Mexico	813	21573	0.227 (0.032, 0.422)	15.425
United States	20247	133178	0.019 (-0.009, 0.047)	397.573
China	1767	7705	0.103 (-0.018, 0.225)	33.245
Japan	354	4380	0.096 (-0.534, 0.725)	1.823
South Korea	1082	21527	0.074 (-0.149, 0.297)	7.250
All Countries (I^2 = 52.6%, 0.3–67.5)			0.005 (-0.007, 0.018)	



Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.4 Association Between Retirement and Diabetes by Country

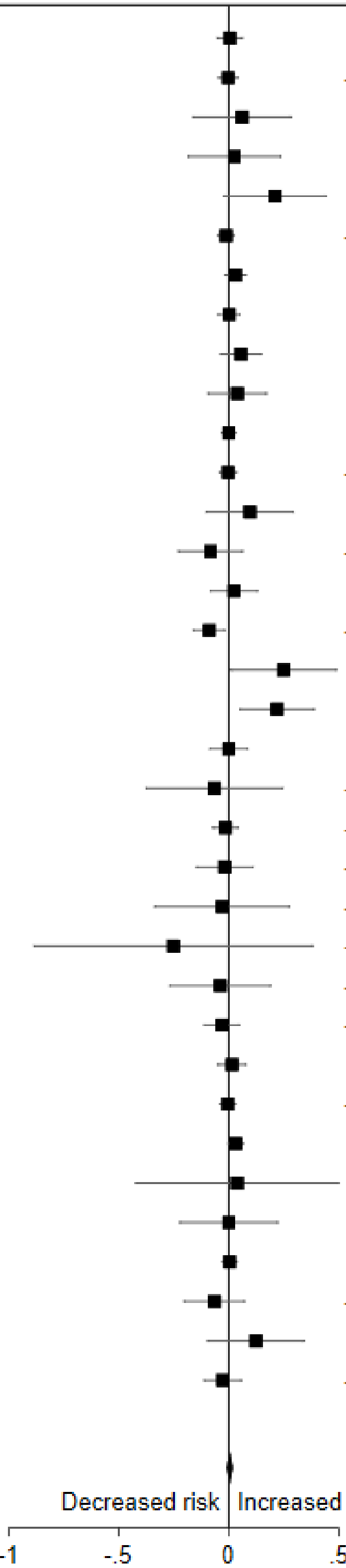


Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.5 Association Between Retirement and Obesity by Country

Obesity

Country	Events	Obs.	Coef. (95% CI)	F-statistic
Austria	986	9254	0.005 (-0.051, 0.061)	119.590
Belgium	1283	13338	-0.003 (-0.045, 0.040)	240.945
Bulgaria	91	704	0.061 (-0.160, 0.282)	12.344
Croatia	341	2581	0.026 (-0.182, 0.234)	15.685
Cyprus	32	238	0.210 (-0.023, 0.442)	15.955
Czech Republic	1732	12478	-0.012 (-0.048, 0.024)	403.225
Denmark	901	10518	0.031 (-0.017, 0.078)	303.598
England	1757	13698	0.002 (-0.047, 0.051)	289.127
Estonia	2092	11580	0.056 (-0.036, 0.147)	53.083
Finland	87	1068	0.039 (-0.093, 0.171)	20.810
France	1299	11878	0.000 (-0.030, 0.030)	443.051
Germany	1103	10754	-0.003 (-0.041, 0.036)	305.571
Greece	523	5948	0.096 (-0.101, 0.293)	10.034
Hungary	267	1806	-0.082 (-0.227, 0.062)	77.739
Israel	456	4349	0.024 (-0.082, 0.131)	53.439
Italy	873	9555	-0.089 (-0.159, -0.019)	95.583
Latvia	71	588	0.250 (0.011, 0.488)	8.548
Lithuania	148	1054	0.219 (0.052, 0.387)	16.381
Luxembourg	223	2313	0.000 (-0.085, 0.085)	52.754
Malta	25	434	-0.065 (-0.374, 0.244)	10.144
Netherlands	452	4844	-0.015 (-0.071, 0.041)	275.840
Poland	714	4418	-0.018 (-0.145, 0.110)	68.089
Portugal	180	1688	-0.030 (-0.335, 0.275)	8.932
Romania	114	1106	-0.251 (-0.883, 0.381)	3.540
Slovakia	42	1288	-0.039 (-0.267, 0.190)	15.934
Slovenia	953	7651	-0.031 (-0.111, 0.049)	78.996
Spain	750	7781	0.015 (-0.048, 0.078)	190.336
Sweden	997	10039	-0.004 (-0.037, 0.029)	254.269
Switzerland	483	7798	0.032 (-0.001, 0.065)	173.629
Costa Rica	75	2566	0.039 (-0.420, 0.497)	1.157
Mexico	737	18361	0.000 (-0.221, 0.222)	13.824
United States	20045	131283	0.004 (-0.027, 0.036)	393.115
China	761	3363	-0.065 (-0.201, 0.072)	10.523
Japan	353	4364	0.123 (-0.094, 0.341)	1.948
South Korea	1072	21322	-0.027 (-0.111, 0.057)	7.417
All Countries (I^2 = 0.0%, 0.0–23.2)			0.004 (-0.009, 0.018)	

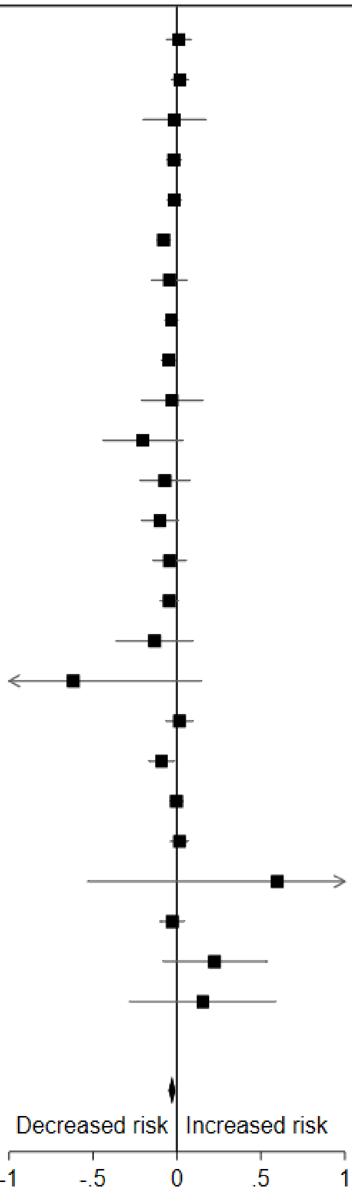


Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.6 Association Between Retirement and Physical Inactivity by Country

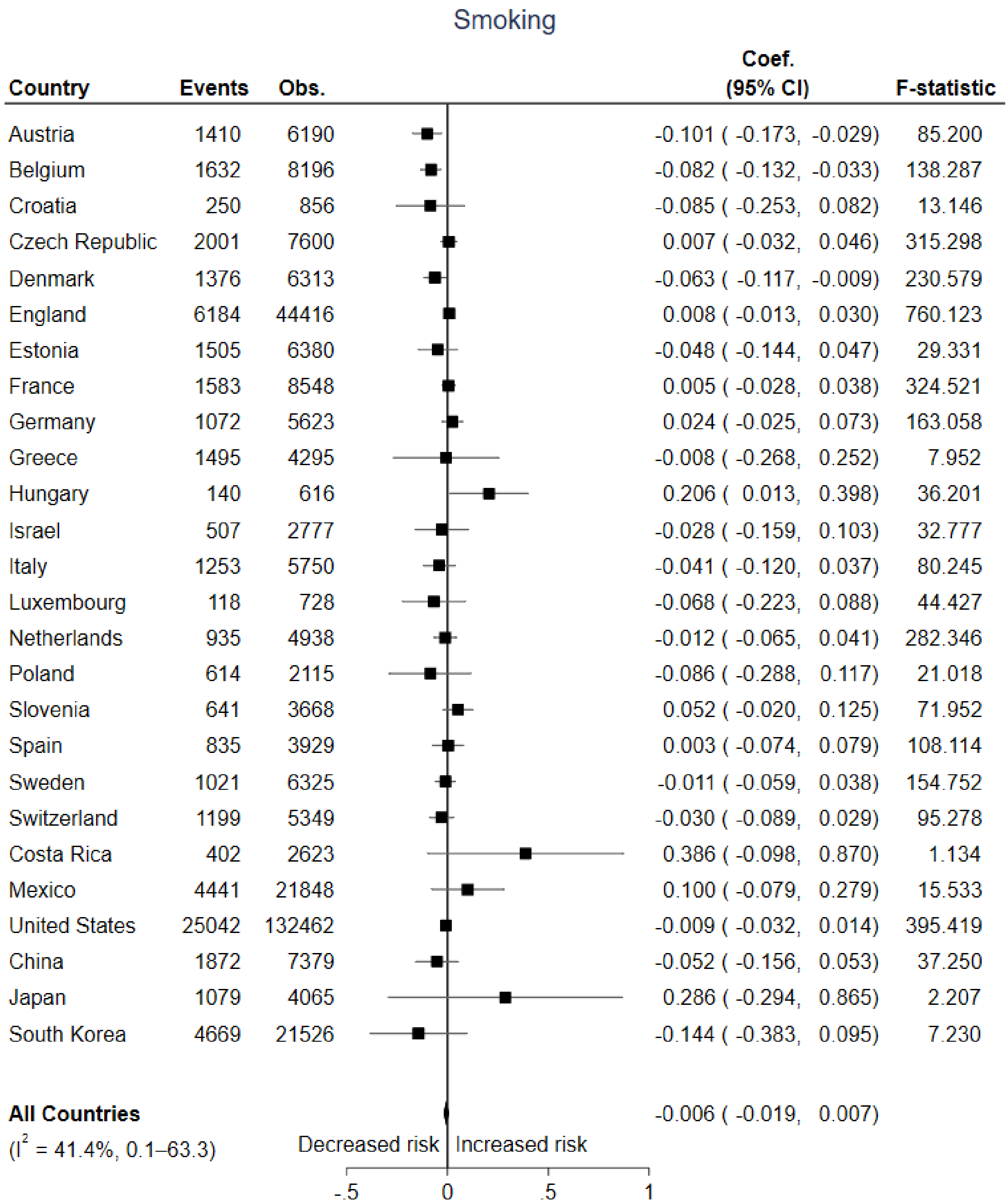
Physical Inactivity

Country	Events	Obs.	Coef. (95% CI)	F-statistic
Austria	751	8039	0.011 (-0.061, 0.082)	98.854
Belgium	912	11333	0.017 (-0.032, 0.067)	200.728
Croatia	67	858	-0.015 (-0.200, 0.169)	13.153
Czech Republic	1310	11029	-0.018 (-0.058, 0.022)	342.196
Denmark	335	9222	-0.016 (-0.059, 0.027)	279.371
England	6183	44647	-0.079 (-0.118, -0.039)	765.056
Estonia	832	9009	-0.044 (-0.151, 0.062)	43.071
France	1009	10725	-0.033 (-0.070, 0.005)	408.387
Germany	538	9061	-0.050 (-0.091, -0.009)	262.337
Greece	535	5403	-0.030 (-0.211, 0.152)	11.173
Hungary	129	616	-0.203 (-0.442, 0.035)	36.201
Israel	479	3782	-0.072 (-0.221, 0.077)	42.785
Italy	1525	7937	-0.102 (-0.211, 0.008)	86.235
Luxembourg	112	1624	-0.044 (-0.144, 0.055)	46.598
Netherlands	207	4938	-0.046 (-0.101, 0.008)	282.346
Poland	495	3070	-0.133 (-0.359, 0.094)	34.830
Portugal	230	1134	-0.617 (-1.382, 0.147)	3.455
Slovenia	434	5438	0.015 (-0.066, 0.096)	68.296
Spain	715	6613	-0.092 (-0.165, -0.019)	162.575
Sweden	364	9150	-0.003 (-0.037, 0.031)	221.057
Switzerland	444	7015	0.015 (-0.035, 0.066)	159.025
Costa Rica	628	1870	0.598 (-0.528, 1.724)	1.379
United States	18948	75515	-0.027 (-0.097, 0.043)	171.736
China	1347	3269	0.224 (-0.084, 0.533)	18.288
South Korea	13590	21527	0.154 (-0.281, 0.589)	7.250
All Countries (I^2 = 27.5%, 0.0–55.6)			-0.030 (-0.049, -0.010)	



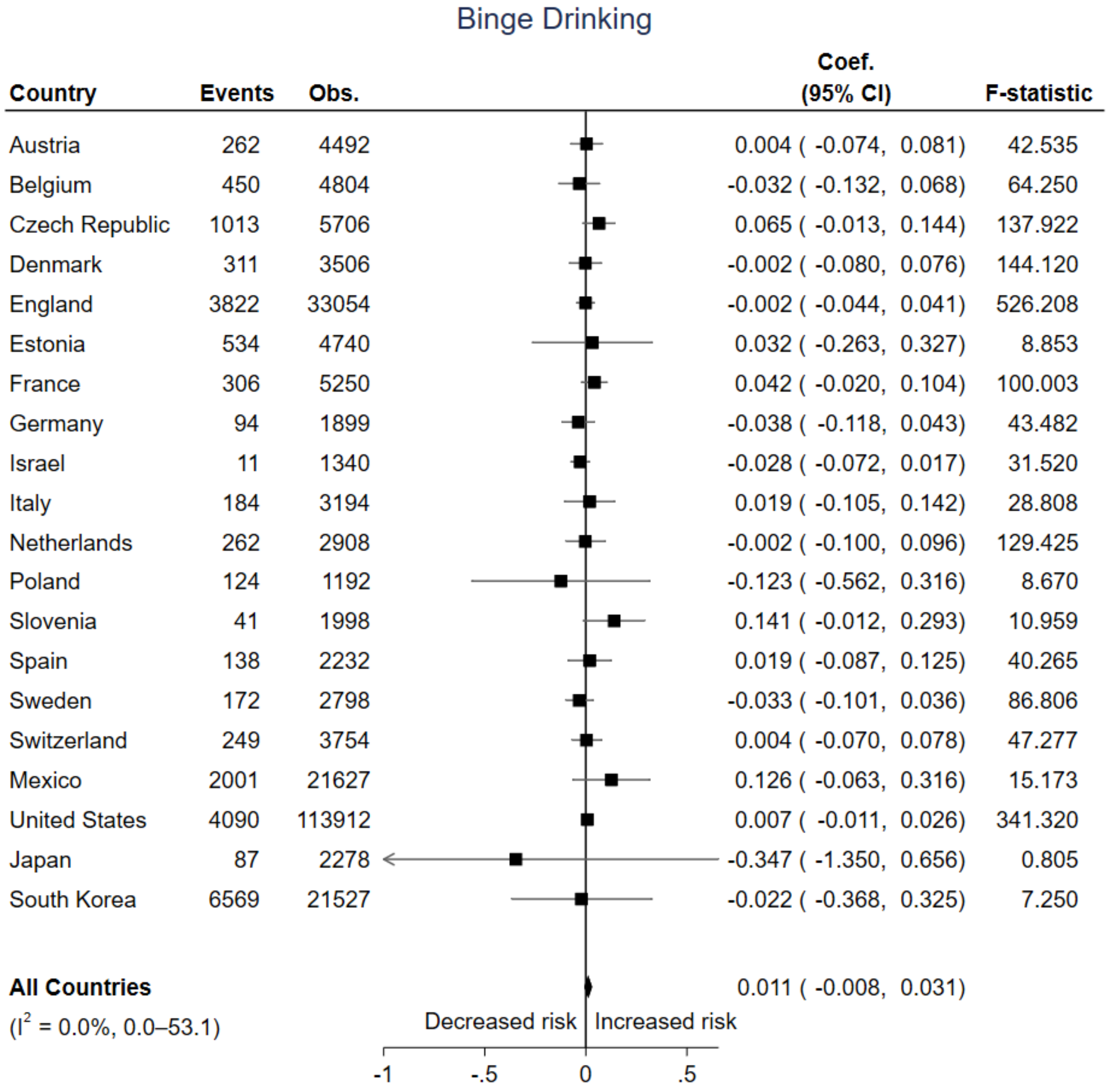
Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.7 Association Between Retirement and Smoking by Country



Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Figure F.8 Association Between Retirement and Binge Drinking by Country



Kleibergen-Paap rk Wald F statistic, I^2 statistic, and its 95% CI were indicated. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals and years.

Table F.2 Models Including Interactions of Regions for the Associations Between Retirement and Outcomes

Outcome	Event	Obs.	Variable	Coef.	95% CI		P-value
Heart disease	47846	395952	Retire	-0.023	-0.033	-0.014	<0.001
			Retire x America	0.006	-0.019	0.030	0.636
			Retire x Asia	0.010	-0.074	0.094	0.811
Stroke	14446	395937	Retire	-0.004	-0.009	0.002	0.163
			Retire x America	0.009	-0.006	0.024	0.252
			Retire x Asia	-0.035	-0.102	0.031	0.294
Hypertension	163273	395974	Retire	0.005	-0.007	0.018	0.408
			Retire x America	0.023	-0.008	0.054	0.148
			Retire x Asia	0.122	0.008	0.235	0.035
Diabetes	54002	395857	Retire	-0.005	-0.014	0.003	0.215
			Retire x America	0.017	-0.008	0.042	0.182
			Retire x Asia	0.028	-0.055	0.111	0.511
Obesity	88523	352008	Retire	0.010	-0.003	0.024	0.138
			Retire x America	-0.008	-0.043	0.027	0.649
			Retire x Asia	0.006	-0.071	0.083	0.877
Physical inactivity	52119	272824	Retire	-0.036	-0.052	-0.020	<0.001
			Retire x America	0.006	-0.066	0.078	0.862
			Retire x Asia	0.019	-0.237	0.274	0.886
Smoking	63276	324519	Retire	-0.010	-0.023	0.003	0.138
			Retire x America	0.004	-0.024	0.031	0.799
			Retire x Asia	-0.058	-0.158	0.043	0.259
Binge drinking	20720	242211	Retire	0.006	-0.020	0.032	0.658
			Retire x America	0.002	-0.031	0.034	0.914
			Retire x Asia	-0.018	-0.354	0.319	0.918

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. European countries were set to the reference group. The number of observations varied across outcomes due to missing values in outcomes.

Table F.3 Models Including Interactions of Low-Middle Income Countries for the Associations Between Retirement and Outcomes

Outcome	Event	Obs.	Variable	Coef.	95% CI		P for interaction
Heart disease	47846	395952	Retire	-0.023	-0.032	-0.013	<0.001
			Retire x LMIC	-0.050	-0.134	0.034	0.242
Stroke	14446	395937	Retire	-0.002	-0.008	0.003	0.412
			Retire x LMIC	-0.025	-0.089	0.039	0.440
Hypertension	163273	395974	Retire	0.002	-0.011	0.014	0.805
			Retire x LMIC	0.127	0.017	0.237	0.023
Diabetes	54002	395857	Retire	-0.003	-0.013	0.006	0.483
			Retire x LMIC	-0.083	-0.168	0.002	0.055
Obesity	88523	352008	Retire	0.005	-0.009	0.019	0.458
			Retire x LMIC	-0.087	-0.249	0.075	0.291
Physical inactivity	52119	272824	Retire	-0.029	-0.048	-0.009	<0.001
			Retire x LMIC	0.139	-0.150	0.427	0.347
Smoking	63276	324519	Retire	-0.006	-0.019	0.006	0.328
			Retire x LMIC	-0.008	-0.099	0.084	0.869
Binge drinking	20720	242211	Retire	0.013	-0.007	0.032	0.199
			Retire x LMIC	0.114	-0.077	0.304	0.242

LMIC stands for low-middle income countries. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. High income countries were set to the reference group. The number of observations varied across outcomes due to missing values in outcomes.

Table F.4 Main Models for the Associations of Retirement with Outcomes

Outcome	Model	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	FE	47846	395952	0.007	0.004	0.010	<0.001		
	FEIV	47846	395952	-0.022	-0.031	-0.012	<0.001	2262.7	0.481
Stroke	FE	14446	395937	0.005	0.003	0.006	<0.001		
	FEIV	14446	395937	-0.002	-0.008	0.004	0.419	2263.0	0.682
Hypertension	FE	163273	395974	0.009	0.0055	0.013	<0.001		
	FEIV	163273	395974	0.005	-0.007	0.018	0.402	2265.0	0.534
Diabetes	FE	54002	395857	0.005	0.002	0.007	0.001		
	FEIV	54002	395857	-0.004	-0.014	0.005	0.378	2264.7	0.416
Obesity	FE	88523	352008	0.002	-0.002	0.006	0.309		
	FEIV	88523	352008	0.004	-0.009	0.018	0.536	1986.2	0.816
Physical inactivity	FE	52119	272824	-0.016	-0.021	-0.010	<0.001		
	FEIV	52119	272824	-0.030	-0.049	-0.010	0.003	1854.4	0.116
Smoking	FE	63276	324519	-0.011	-0.015	-0.008	<0.001		
	FEIV	63276	324519	-0.006	-0.019	0.007	0.351	1592.1	0.057
Binge drinking	FE	20720	242211	-0.004	-0.008	-0.001	0.016		
	FEIV	20720	242211	0.011	-0.008	0.031	0.253	788.3	0.424

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.5 Models Including Interactions of Sex for the Associations Between Retirement and Outcomes

Outcome	Event	Obs.	Variable	Coef.	95% CI		P-value
Heart disease	47846	395952	Retire	-0.029	-0.040	-0.018	<0.001
			Retire x Men	0.002	-0.017	0.022	0.798
Stroke	14446	395937	Retire	-0.003	-0.010	0.004	0.426
			Retire x Men	-0.004	-0.016	0.008	0.531
Hypertension	163273	395974	Retire	0.001	-0.015	0.017	0.914
			Retire x Men	0.004	-0.021	0.029	0.762
Diabetes	54002	395857	Retire	-0.010	-0.021	0.002	0.098
			Retire x Men	0.006	-0.013	0.025	0.516
Obesity	88523	352008	Retire	0.010	-0.008	0.028	0.265
			Retire x Men	-0.015	-0.043	0.012	0.278
Physical inactivity	52119	272824	Retire	-0.040	-0.064	-0.015	0.001
			Retire x Men	0.020	-0.018	0.058	0.306
Smoking	63276	324519	Retire	-0.024	-0.038	-0.009	0.002
			Retire x Men	0.048	0.022	0.074	<0.001
Binge drinking	20720	242211	Retire	-0.002	-0.024	0.020	0.877
			Retire x Men	0.024	-0.015	0.063	0.221

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. Women were set to the reference group. The number of observations varied across outcomes due to missing values in outcomes.

Table F.6 Models Including Interactions of Education for the Associations Between Retirement and Outcomes

Outcome	Event	Obs.	Variable	Coef.	95% CI		P-value
Heart disease	46295	370766	Retire	-0.024	-0.037	-0.010	<0.001
			Retire x Low	-0.005	-0.026	0.016	0.636
			Retire x High	-0.004	-0.029	0.021	0.764
Stroke	13648	370755	Retire	0.007	0.000	0.015	0.066
			Retire x Low	-0.010	-0.024	0.004	0.159
			Retire x High	-0.019	-0.034	-0.005	0.010
Hypertension	155748	370788	Retire	0.007	-0.010	0.024	0.405
			Retire x Low	0.011	-0.018	0.040	0.441
			Retire x High	-0.003	-0.035	0.030	0.879
Diabetes	50867	370671	Retire	-0.0064	-0.019	0.006	0.312
			Retire x Low	0.012	-0.010	0.034	0.294
			Retire x High	0.010	-0.013	0.033	0.401
Obesity	87904	329546	Retire	0.017	-0.002	0.036	0.082
			Retire x Low	-0.003	-0.036	0.031	0.879
			Retire x High	-0.045	-0.079	-0.012	0.008
Physical inactivity	38013	247837	Retire	-0.028	-0.054	-0.002	0.033
			Retire x Low	0.013	-0.030	0.055	0.564
			Retire x High	-0.018	-0.061	0.024	0.402
Smoking	58079	299372	Retire	-0.010	-0.027	0.006	0.229
			Retire x Low	-0.009	-0.038	0.020	0.545
			Retire x High	-0.004	-0.034	0.027	0.819
Binge drinking	13892	218016	Retire	0.004	-0.017	0.026	0.682
			Retire x Low	-0.012	-0.055	0.032	0.597
			Retire x High	0.018	-0.031	0.067	0.466

Educational attainment was asked in the interview, and the harmonized datasets categorized it into three groups based on the 1997 International Standard Classification of Education codes: low=less than upper secondary education; middle=upper secondary and vocational training; high=tertiary education. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. People with middle educational levels were set to the reference group. The number of observations varied across outcomes due to missing values in outcomes.

Table F.7 Models Including Interactions of Job Type for the Associations Between Retirement and Outcomes

Outcome	Event	Obs.	Variable	Coef.	95% CI		P-value
Heart disease	27291	258388	Retire	-0.030	-0.048	-0.012	0.001
			Retire x Physical	0.019	-0.005	0.044	0.114
Stroke	6973	258372	Retire	-0.003	-0.014	0.007	0.539
			Retire x Physical	0.003	-0.012	0.018	0.663
Hypertension	98956	258404	Retire	0.017	-0.007	0.040	0.161
			Retire x Physical	-0.005	-0.037	0.026	0.750
Diabetes	30521	258388	Retire	0.010	-0.008	0.027	0.269
			Retire x Physical	-0.006	-0.029	0.018	0.647
Obesity	57124	234205	Retire	-0.031	-0.055	-0.007	0.010
			Retire x Physical	0.055	0.022	0.087	0.001
Physical inactivity	35654	194225	Retire	-0.042	-0.076	-0.009	0.014
			Retire x Physical	0.039	-0.008	0.086	0.101
Smoking	42320	223630	Retire	-0.015	-0.036	0.006	0.165
			Retire x Physical	0.022	-0.007	0.051	0.144
Binge drinking	15366	174681	Retire	0.019	-0.009	0.047	0.179
			Retire x Physical	-0.023	-0.062	0.017	0.266

The surveys ask participants in paid work to choose from four options about how much they agree that their job is physically demanding: ‘strongly disagree’, ‘disagree’, ‘agree’, or ‘strongly agree’. We dichotomised their responses and considered those who answered ‘agree’ or ‘strongly agree’ at least once in the interview as having the experience of physical labor, and otherwise as having engaged in non-physical labor. We excluded those who did not engage in paid work during the study period from this subgroup analysis and compared retirees with workers within the job type categories. The question was not asked in CRELES, MHAS, and CHARLS.

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. People who have not experienced physical labor were set to the reference group. The number of observations varied across outcomes due to missing values in outcomes.

Table F.8 Associations of Full Retirement with Outcomes

Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	44051	367935	-0.021	-0.033	-0.010	<0.001	1711.2	0.427
Stroke	13498	367921	-0.001	-0.008	0.007	0.834	1711.5	0.920
Hypertension	150832	367956	0.006	-0.009	0.021	0.447	1713.4	0.676
Diabetes	50186	367840	-0.002	-0.013	0.010	0.756	1713.0	0.745
Obesity	81674	325276	0.007	-0.010	0.024	0.414	1447.6	0.910
Physical inactivity	49670	249571	-0.025	-0.050	0.0003	0.053	1301.3	0.055
Smoking	59139	299763	-0.003	-0.019	0.013	0.741	1147.3	0.223
Binge drinking	19694	223520	0.014	-0.010	0.038	0.259	586.8	0.215

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.9 Associations of Retirement with Outcomes for Participants Aged 52-68

Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	38603	333121	-0.020	-0.031	-0.009	<0.001	1377.1	0.839
Stroke	11469	333110	-0.003	-0.010	0.004	0.374	1377.5	0.740
Hypertension	135808	333142	-0.015	-0.030	0.001	0.067	1379.1	0.530
Diabetes	44581	333036	-0.013	-0.024	-0.002	0.023	1378.6	0.196
Obesity	74992	296007	0.011	-0.006	0.028	0.198	1204.1	0.777
Physical inactivity	42809	226667	-0.033	-0.058	-0.008	0.011	1079.6	0.220
Smoking	53745	271785	0.003	-0.014	0.020	0.723	925.7	0.032
Binge drinking	17529	203917	0.022	-0.005	0.049	0.109	469.1	0.792

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.10 Models Excluding Weak IVs for the Associations of Retirement with Outcomes

Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	45406	357497	-0.024	-0.034	-0.015	<0.001	2454.7	0.399
Stroke	13280	357485	-0.002	-0.008	0.004	0.552	2455.1	0.647
Hypertension	150587	357519	0.006	-0.007	0.018	0.371	2457.0	0.833
Diabetes	49029	357416	-0.003	-0.012	0.006	0.481	2456.6	0.493
Obesity	85179	313992	0.005	-0.009	0.019	0.449	2191.7	0.794
Physical inactivity	37136	242890	-0.038	-0.055	-0.020	<0.001	2097.8	0.051
Smoking	55631	292006	-0.009	-0.021	0.003	0.164	1774.8	0.154
Binge drinking	14064	218406	0.005	-0.013	0.022	0.604	901.7	0.086

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.11 Models Excluding the United States for the Associations of Retirement with Outcomes

Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	27599	262774	-0.020	-0.029	-0.010	<0.001	2102.9	0.470
Stroke	8399	262759	-0.005	-0.011	0.001	0.091	2103.4	0.971
Hypertension	101117	262796	0.011	-0.003	0.024	0.116	2105.7	0.657
Diabetes	31314	262679	-0.008	-0.017	0.001	0.096	2105.7	0.321
Obesity	44999	220725	0.007	-0.007	0.021	0.316	1871.4	0.695
Physical inactivity	33171	197309	-0.029	-0.047	-0.010	0.002	1843.3	0.638
Smoking	38234	192057	-0.008	-0.023	0.007	0.318	1381.1	0.039
Binge drinking	16630	128299	0.015	-0.019	0.049	0.383	498.5	0.234

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.12 Models Using Multiple Imputation for the Associations of Retirement with Outcomes

Outcome	Event	Obs.	Coef.	95% CI		P-value
Heart disease	48511	402077	-0.022	-0.031	-0.012	<0.001
Stroke	14692	402077	-0.002	-0.008	0.003	0.409
Hypertension	165454	402077	0.004	-0.009	0.017	0.528
Diabetes	54787	402077	-0.005	-0.015	0.004	0.274
Obesity	102156	402077	-0.010	-0.029	0.008	0.270
Physical inactivity	61052	316948	-0.024	-0.043	-0.004	0.018
Smoking	80091	402077	-0.014	-0.031	0.003	0.111
Binge drinking	21869	253546	0.012	-0.008	0.033	0.225

We adopted multiple imputation using the algorithm of expectation-maximisation with bootstrapping and created ten imputed datasets. Our imputation model had a hierarchical structure nesting observations within an individual with a linear time trend. Assuming missing at random, sex, age, marital status, working status, retirement status, eight outcome variables, and country were used to predict missing values. The following outcomes not obtained in all or most waves were not imputed: physical inactivity in all waves of MHAS and JSTAR, the first cohort (waves 1–2) of CRELES, and waves 1–6 of HRS; binge drinking in waves 1 and 6–8 of SHARE, wave 1 of ELSA, waves 1–2 of HRS, and all waves of CRELES and CHARLS. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years.

Table F.13 Poisson Fixed-Effects Instrumental Variable Models for the Associations of Retirement with Outcomes

Outcome	Event	Obs.	RR	95% CI		P-value
Heart disease	47846	395952	0.89	0.81	0.98	0.017
Stroke	14446	395937	1.24	0.98	1.58	0.070
Hypertension	163273	395974	0.99	0.94	1.03	0.565
Diabetes	54002	395857	0.95	0.87	1.03	0.241
Obesity	88523	352008	1.03	0.97	1.10	0.348
Physical inactivity	52119	272824	0.87	0.77	0.97	0.015
Smoking	63276	324519	1.03	0.95	1.13	0.466
Binge drinking	20720	242211	1.10	0.82	1.49	0.513

RR denotes risk ratio. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. Standard errors were calculated using clustered bootstrapping at the individual level.

Table F.14 Models Excluding Participants Who Retired Within Two Years

Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	44028	367848	-0.029	-0.040	-0.017	<0.001	1827.5	0.414
Stroke	13343	367837	-0.002	-0.009	0.006	0.673	1827.3	0.408
Hypertension	150800	367867	0.003	-0.012	0.018	0.709	1829.7	0.785
Diabetes	49952	367762	-0.006	-0.017	0.005	0.263	1828.5	0.416
Obesity	81786	326484	0.004	-0.013	0.020	0.662	1612.4	0.833
Physical inactivity	48347	249736	-0.035	-0.058	-0.011	0.004	1461.4	0.032
Smoking	58405	299035	-0.007	-0.022	0.008	0.369	1237.0	0.106
Binge drinking	19381	223000	0.004	-0.019	0.027	0.743	609.1	0.502

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

Table F.15 FEIV Models Including Interactions Between IVs and Marital Status

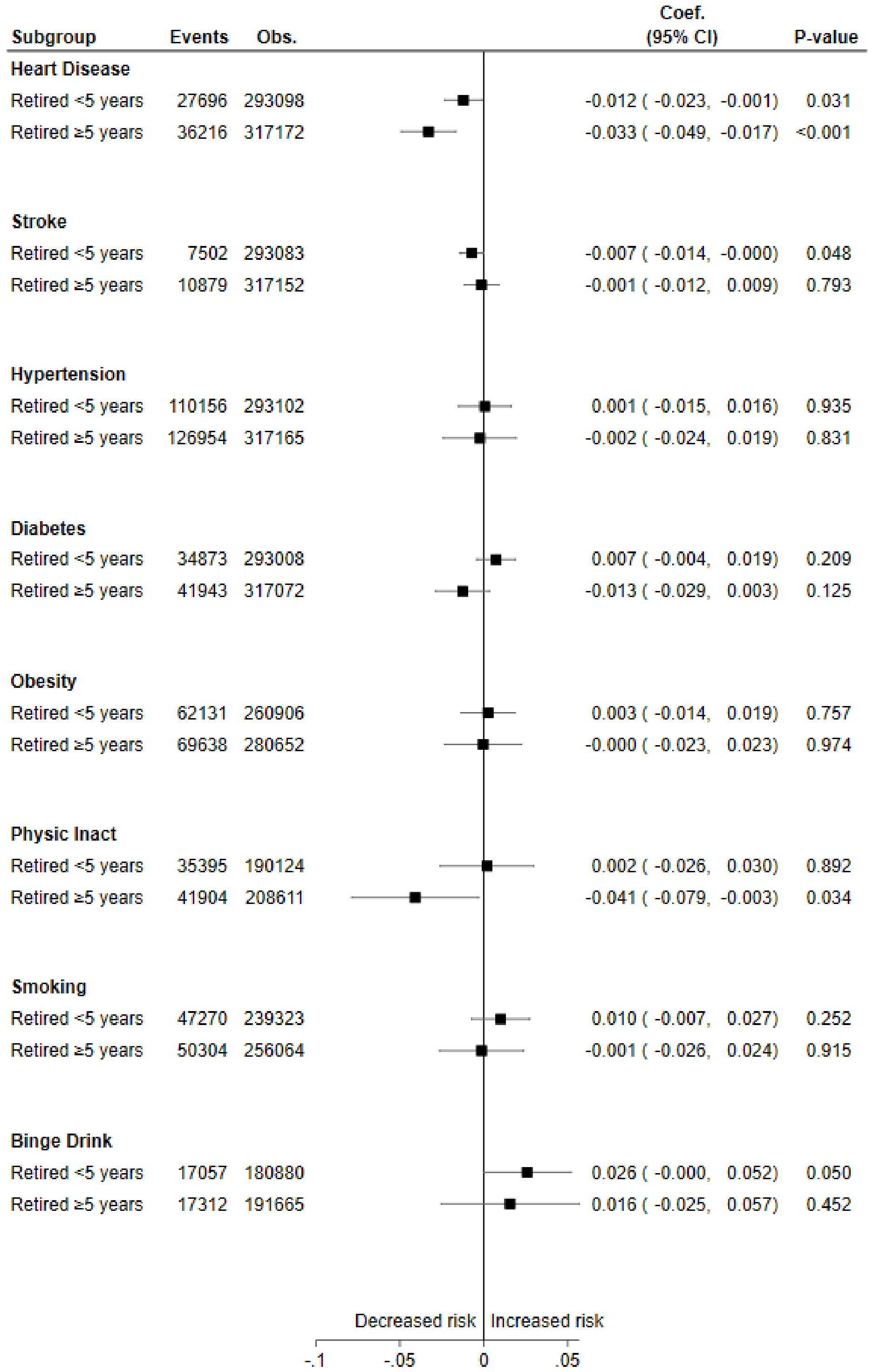
Outcome	Event	Obs.	Coef.	95% CI		P-value	F statistic [a]	P for J statistic [b]
Heart disease	47846	395952	-0.022	-0.031	-0.013	0.000	1699.9	0.481
Stroke	14446	395937	-0.003	-0.009	0.003	0.307	1700.1	0.682
Hypertension	163273	395974	0.006	-0.007	0.019	0.340	1701.5	0.534
Diabetes	54002	395857	-0.005	-0.014	0.004	0.304	1701.3	0.416
Obesity	88523	352008	0.005	-0.009	0.019	0.497	1492.3	0.816
Physical inactivity	52119	272824	-0.030	-0.049	-0.010	0.003	1394.6	0.116
Smoking	63276	324519	-0.004	-0.017	0.009	0.561	1196.0	0.057
Binge drinking	20720	242211	0.012	-0.008	0.031	0.249	591.3	0.424

All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes.

[a] Kleibergen-Paap rk Wald F statistic for a weak identification test.

[b] Hansen's J statistic for an over-identification test of instruments of early retirement age and official retirement age.

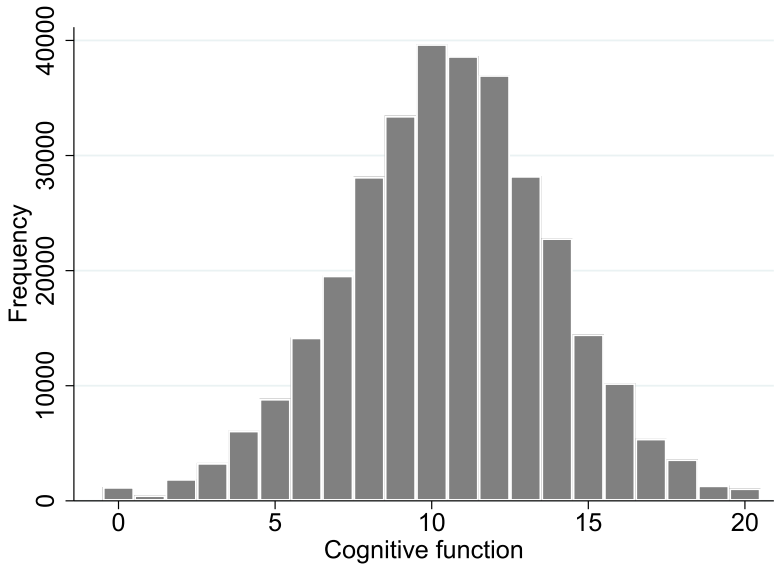
Figure F.9 Subgroup Analysis by Retirement Timing for the Association of Retirement with Cardiovascular Diseases and Risk Factors



Physic Inact. denotes physical inactivity. All models were adjusted for age, age squared, marital status, and fixed-effects of individuals, countries, and years. The number of observations varied across outcomes due to missing values in outcomes. JSTAR does not provide information on retirement timing.

G. Supplementary Materials for Chapter 2

Figure G.1 Distribution of Cognitive Function Score



Data included only surveys using a 10-word list.

Figure G.2 Distribution of Physical Function Score

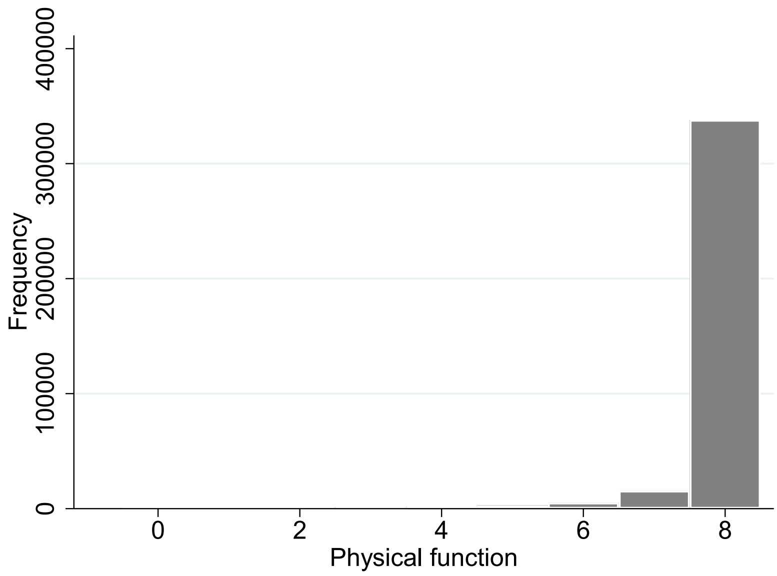


Figure G.3 Distribution of Self-Rated Health Score

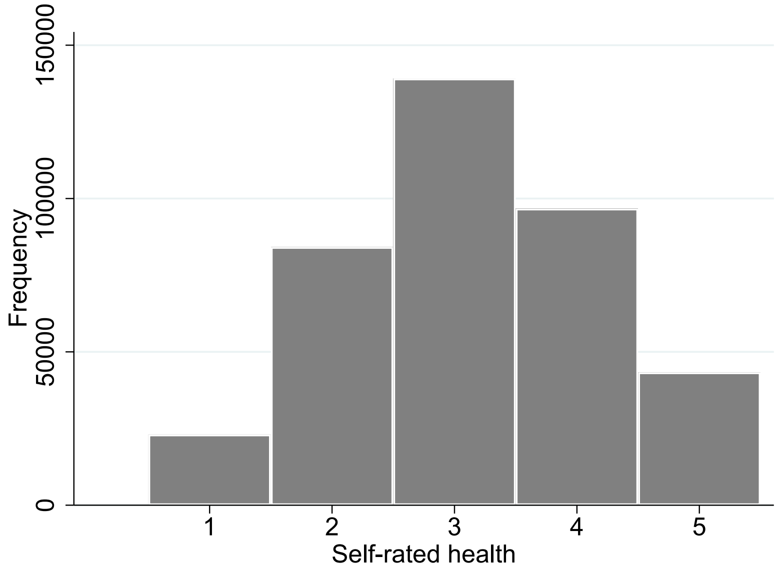


Table G.1 Characteristics Comparison Between Followed Up and Lost to Follow-Up Participants

Characteristics in the previous interview	Lost to follow-up		Followed up		Standardized difference
	Mean (Obs.)	SD (%)	Mean (Obs.)	SD (%)	
Retired	0.487	0.500	0.441	0.497	0.092
Outcome variables					
Health status					
Cognitive function (z-score)	-0.011	1.019	0.063	0.972	-0.074
Physical independence	0.918	0.275	0.925	0.264	-0.026
Self-rated health (z-score)	0.003	1.007	0.112	0.961	-0.110
Health behavior as risk factors					
Physical inactivity	0.180	0.384	0.190	0.392	-0.026
Smoking	0.226	0.418	0.196	0.397	0.072
Binge drinking	0.085	0.278	0.086	0.280	-0.003
Covariates					
Age	61.28	5.406	60.44	5.509	0.154
Married	0.783	0.412	0.783	0.413	0.001
Potential effect of heterogeneity					
Men	0.525	0.499	0.494	0.500	0.061
Education					0.098
Low	(13,700)	(31.7)	(79,318)	(27.4)	
Middle	(19,175)	(44.3)	(139,831)	(48.2)	
High	(10,407)	(24.0)	(70,693)	(24.4)	
Physically demanding job	0.537	0.499	0.559	0.497	-0.043
Low control job	0.354	0.478	0.342	0.474	0.025

The scores of cognitive function and self-rated health are standardized. In general, a standardized difference less than 0.1 indicates a well balance between the two groups.

Table G.2 First Stage Estimates of Adjusted FEIV Models

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Age	0.017***	0.018***	0.016***	0.021***	0.015***	0.013***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Age^2	0.027***	0.027***	0.026***	0.029***	0.027***	0.023***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Married	0.007**	0.003	0.005*	-0.001	0.005	0.004
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
ERA	0.086***	0.083***	0.084***	0.087***	0.083***	0.064***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
ERA x Age	-0.006***	-0.007***	-0.006***	-0.008***	-0.008***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ERA x Age^2	-0.005***	-0.004***	-0.004***	-0.002**	-0.002**	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ORA	0.180***	0.182***	0.175***	0.180***	0.165***	0.123***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
ORA x Age	-0.010***	-0.009***	-0.009***	-0.004***	-0.009***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
ORA x Age^2	-0.035***	-0.036***	-0.035***	-0.042***	-0.034***	-0.033***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	377,276	362,973	384,631	272,824	324,519	242,211

Age squared was divided by 10 for ease of interpretation. All regressions are adjusted for fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table G.3 FEIV Models with an Interaction of Sex

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.096***	0.034***	0.184***	-0.040***	-0.024***	-0.002
	(0.026)	(0.008)	(0.024)	(0.013)	(0.008)	(0.011)
Retirement x Men	-0.088**	-0.015	-0.071*	0.020	0.048***	0.024
	(0.041)	(0.012)	(0.037)	(0.019)	(0.013)	(0.020)
Observations	377,276	362,973	384,631	272,824	324,519	242,211
Kleibergen-Paap F	505.009	499.381	497.442	402.296	357.436	208.600
Hansen J	0.850	1.149	3.897**	4.070**	3.175*	0.417

All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and sex, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.4 FEIV Models with Interactions of Education

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.050*	0.020**	0.153***	-0.028**	-0.010	0.004
	(0.027)	(0.008)	(0.024)	(0.013)	(0.009)	(0.011)
Retirement x Low education	-0.025	-0.002	-0.043	0.013	-0.009	-0.012
	(0.045)	(0.015)	(0.043)	(0.022)	(0.015)	(0.022)
Retirement x High education	0.032	0.009	0.025	-0.018	-0.004	0.018
	(0.054)	(0.013)	(0.047)	(0.022)	(0.016)	(0.025)
Observations	353,219	337,831	360,025	247,837	299,372	218,016
Kleibergen-Paap F	147.951	153.754	144.239	132.078	97.432	48.761
Hansen J	2.628	2.729*	5.220**	10.999***	9.587***	4.423**

All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and education, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.5 FEIV Models with an Interaction of a Physically Demanding Job

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.068*	0.036***	0.154***	-0.042**	-0.015	0.019
	(0.037)	(0.009)	(0.032)	(0.017)	(0.011)	(0.014)
Retirement x Physically demanding	-0.062	-0.015	-0.025	0.039	0.022	-0.023
	(0.050)	(0.014)	(0.045)	(0.024)	(0.015)	(0.020)
Observations	247,236	233,908	254,372	194,225	223,630	174,681
Kleibergen-Paap F	361.632	328.579	375.916	316.143	440.371	257.944
Hansen J	0.308	1.245	6.694**	1.243	3.286*	1.359

All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and engagement in a physically demanding job, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.6 FEIV Models with an Interaction of a Low Control Job

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.054	0.003	0.149***	-0.044***	-0.017	0.009
	(0.037)	(0.009)	(0.033)	(0.015)	(0.012)	(0.018)
Retirement x Low control	-0.030	0.022	0.072	0.025	0.039**	-0.022
	(0.056)	(0.014)	(0.051)	(0.023)	(0.018)	(0.028)
Observations	161,552	161,559	161,606	144,311	131,504	97,343
Kleibergen-Paap F	261.945	268.037	254.415	267.031	177.343	103.537
Hansen J	1.125	0.341	13.919***	1.130	1.980	7.049***

Note: FEIV denotes fixed effect with instrumental variable. All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and engagement in a low control job, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.7 FEIV Models with Interactions of Region

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.055***	0.019***	0.138***	-0.036***	-0.010	0.006
	(0.020)	(0.005)	(0.019)	(0.008)	(0.007)	(0.013)
Retirement x America	-0.022	0.025	-0.021	0.006	0.004	0.002
	(0.051)	(0.019)	(0.043)	(0.037)	(0.014)	(0.017)
Retirement x Asia	0.129	-0.034	-0.184	0.019	-0.058	-0.018
	(0.194)	(0.058)	(0.300)	(0.130)	(0.051)	(0.172)
Observations	377,276	362,973	384,631	272,824	324,519	242,211
Kleibergen-Paap F	8.540	9.443	4.151	6.001	10.023	2.706
Hansen J	5.166**	0.963	10.413***	6.725**	2.719*	0.960

All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and regions, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.8 FEIV Models with an Interaction of Country Income Level

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.047**	0.030***	0.151***	-0.029***	-0.006	0.013
	(0.021)	(0.006)	(0.019)	(0.010)	(0.007)	(0.010)
Retirement x LMIC	-0.198	-0.043	-0.258	0.139	-0.008	0.114
	(0.164)	(0.059)	(0.236)	(0.147)	(0.047)	(0.097)
Observations	377,276	362,973	384,631	272,824	324,519	242,211
Kleibergen-Paap F	16.236	16.814	9.357	5.596	17.492	7.590
Hansen J	1.606	0.413	2.850*	3.977**	7.636***	0.493

LMIC denotes low-middle income countries. All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and country income level, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.9 FEIV Models with an Interaction of Aged Society

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.032	0.043**	0.096**	0.009	0.004	0.002
	(0.049)	(0.018)	(0.042)	(0.044)	(0.013)	(0.012)
Retirement x Aged society	0.018	-0.022	0.042	-0.046	-0.012	0.010
	(0.053)	(0.019)	(0.046)	(0.045)	(0.015)	(0.018)
Observations	377,276	362,973	384,631	272,824	324,519	242,211
Kleibergen-Paap F	184.996	142.173	191.637	69.539	186.734	152.033
Hansen J	0.494	2.260	5.451**	2.057	3.310*	1.555

All regressions are adjusted for covariates (age, age squared, and marital status), interactions between covariates and aged society, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.10 FEIV Models for the Effect of Full-Retirement on Outcomes

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Full retirement	0.037	0.035***	0.150***	-0.025*	-0.003	0.014
	(0.025)	(0.007)	(0.023)	(0.013)	(0.008)	(0.012)
Observations	349,739	336,718	356,752	249,571	299,763	223,520
Kleibergen-Paap F	1681.882	1688.270	1634.389	1302.014	1147.875	586.978
Hansen J	0.090	0.048	3.954**	3.671*	1.484	1.540

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.11 FEIV Models According to Pre-Retirement Employment Status

	Cognitive function			Physical independence			Self-rated health		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Retirement	0.042*	0.059**	0.056	0.020***	0.019***	0.030***	0.135***	0.151***	0.121***
	(0.024)	(0.024)	(0.035)	(0.007)	(0.007)	(0.010)	(0.022)	(0.022)	(0.032)
Observations	292,796	213,426	121,608	280,106	204,848	116,951	299,649	217,932	125,624
Kleibergen-Paap F	1443.132	1392.887	715.787	1424.532	1375.721	702.718	1409.601	1355.670	698.016
Hansen J	0.000	0.002	0.288	0.546	1.260	0.187	3.716*	3.822*	4.519**
	Physical inactivity			Smoking			Binge drinking		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Retirement	-0.017	-0.025**	-0.014	-0.002	-0.001	-0.000	0.012	0.012	0.011
	(0.011)	(0.012)	(0.017)	(0.008)	(0.008)	(0.011)	(0.011)	(0.010)	(0.015)
Observations	209,740	155,283	86,794	261,011	185,434	104,796	200,446	140,680	80,202
Kleibergen-Paap F	1188.162	1113.025	565.052	1096.946	1057.489	584.311	620.373	610.673	336.670
Hansen J	0.681	4.346**	0.503	3.047*	1.879	6.004**	0.021	0.015	0.329

Model 1 restricted the participants to those who answered that they were in paid work at least once in the interviews; Model 2 additionally excluded those who were self-employed from Model1; Model 3 additionally excluded those who experienced a part-time job from Model 2. All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.12 FEIV Models for People Aged 52-68

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	-0.004	0.016**	0.154***	-0.033**	0.003	0.022
	(0.026)	(0.008)	(0.024)	(0.013)	(0.009)	(0.014)
Observations	316,949	303,868	323,716	226,667	271,785	203,917
Kleibergen-Paap F	1343.216	1345.309	1321.374	1080.231	926.187	469.269
Hansen J	0.132	0.025	0.992	1.506	4.614**	0.069

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.13 Country-by-Country FEIV Models for the Effect of Retirement on Cognitive Function

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	9,028	0.148 (0.091)	118.716
Belgium	13,415	0.071 (0.069)	239.358
Bulgaria	752	0.934 (0.357)***	11.900
Croatia	2,596	0.081 (0.334)	16.431
Cyprus	248	-1.097 (0.295)***	15.848
Czech Republic	12,565	-0.022 (0.051)	411.344
Denmark	10,614	0.129 (0.084)	308.888
England	42,988	0.025 (0.048)	729.130
Estonia	10,999	0.159 (0.142)	48.371
Finland	1,058	0.314 (0.183)*	21.825
France	11,874	0.015 (0.051)	440.918
Germany	10,745	0.024 (0.064)	306.879
Greece	5,971	-0.129 (0.339)	9.785
Hungary	1,802	0.059 (0.197)	74.571
Israel	4,468	0.231 (0.183)	51.575
Italy	9,477	-0.027 (0.117)	94.900
Latvia	596	0.026 (0.330)	8.082
Lithuania	1,054	-0.896 (0.303)***	16.359
Luxembourg	2,240	0.031 (0.168)	49.293
Malta	466	-0.411 (0.387)	10.630
Netherlands	4,902	0.226 (0.102)**	278.824
Poland	4,505	0.213 (0.188)	68.037
Portugal	1,634	0.339 (0.510)	7.044
Romania	1,120	0.024 (0.820)	3.568
Slovakia	1,328	0.232 (0.241)	15.314
Slovenia	7,561	0.174 (0.131)	77.866
Spain	7,913	0.050 (0.087)	195.794
Sweden	10,168	0.110 (0.064)*	259.288
Switzerland	7,833	-0.008 (0.067)	177.932
Costa Rica	2,577	-1.686 (1.764)	0.717
Mexico	20,058	-0.070 (0.287)	15.625
United States	125,283	0.020 (0.046)	378.917
China	7,044	-0.539 (0.203)***	29.214
Japan	1,760	0.311 (1.012)	2.718
South Korea	20,634	0.537 (0.420)	7.818

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.14 Country-by-Country FEIV Models for the Effect of Retirement on Physical Independence

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	9,402	0.010 (0.021)	124.056
Belgium	13,481	0.031 (0.018)*	242.973
Bulgaria	738	-0.244 (0.085)***	12.543
Croatia	2,620	0.049 (0.085)	16.425
Cyprus	244	0.093 (0.051)*	16.346
Czech Republic	12,866	-0.005 (0.013)	412.966

Denmark	10,670	0.017 (0.017)	310.583
England	44,684	0.029 (0.014)**	765.277
Estonia	11,707	0.126 (0.040)***	53.531
Finland	1,094	-0.093 (0.045)**	21.660
France	12,034	0.004 (0.011)	447.183
Germany	10,836	0.035 (0.017)**	307.756
Greece	5,991	0.004 (0.057)	10.021
Hungary	1,814	0.105 (0.051)**	79.989
Israel	4,701	-0.028 (0.044)	55.968
Italy	9,645	0.011 (0.023)	97.435
Latvia	600	-0.061 (0.078)	7.702
Lithuania	1,046	0.032 (0.063)	16.707
Luxembourg	2,342	-0.047 (0.026)*	54.277
Malta	472	0.031 (0.089)	10.549
Netherlands	4,937	-0.044 (0.025)*	282.648
Poland	4,549	0.043 (0.055)	67.549
Portugal	1,755	-0.024 (0.135)	7.667
Romania	1,092	-0.098 (0.267)	3.560
Slovakia	1,308	0.179 (0.079)**	14.457
Slovenia	7,767	0.032 (0.027)	83.140
Spain	8,160	-0.020 (0.018)	202.179
Sweden	10,230	-0.003 (0.013)	260.822
Switzerland	7,880	0.007 (0.012)	179.354
Costa Rica	2,599	0.545 (0.453)	1.142
Mexico	18,245	-0.029 (0.087)	14.812
United States	105,591	0.043 (0.018)**	298.573
China	7,963	-0.064 (0.071)	35.715
Japan	2,384	0.285 (0.155)*	5.565
South Korea	21,526	0.126 (0.102)	7.247

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.15 Country-by-Country FEIV Models for the Effect of Retirement on Self-Rated Health

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	9,416	0.346 (0.088)***	123.818
Belgium	13,502	0.234 (0.068)***	242.790
Bulgaria	754	-0.152 (0.251)	12.556
Croatia	2,634	0.138 (0.303)	16.413
Cyprus	248	0.263 (0.276)	15.848
Czech Republic	12,879	0.091 (0.047)*	412.274
Denmark	10,678	0.236 (0.077)***	310.363
England	37,816	0.028 (0.045)	604.670
Estonia	11,747	0.345 (0.123)***	54.254
Finland	1,098	0.619 (0.213)***	21.663
France	12,051	0.093 (0.050)*	447.339
Germany	10,880	0.311 (0.061)***	308.773
Greece	5,990	-0.139 (0.281)	9.981
Hungary	1,838	0.306 (0.186)	78.811
Israel	4,688	0.050 (0.151)	55.869
Italy	9,666	0.016 (0.113)	97.374
Latvia	606	-0.016 (0.289)	8.361

Lithuania	1,056	-0.193 (0.294)	16.378
Luxembourg	2,355	-0.087 (0.155)	55.232
Malta	478	0.463 (0.340)	10.548
Netherlands	4,938	0.037 (0.094)	282.518
Poland	4,562	-0.057 (0.175)	69.196
Portugal	1,753	0.389 (0.451)	7.632
Romania	1,120	0.064 (0.778)	3.568
Slovakia	1,330	0.091 (0.225)	15.383
Slovenia	7,783	0.320 (0.128)**	81.565
Spain	8,163	0.088 (0.086)	201.757
Sweden	10,236	0.065 (0.057)	260.253
Switzerland	7,888	0.151 (0.065)**	179.021
Costa Rica	2,625	0.267 (1.013)	1.133
Mexico	20,093	0.018 (0.310)	15.746
United States	133,109	0.127 (0.038)***	397.264
China	4,701	-0.342 (0.328)	10.867
Japan	4,423	-0.899 (1.199)	1.816
South Korea	21,527	-0.303 (0.419)	7.252

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.16 Country-by-Country FEIV Models for the Effect of Retirement on physical inactivity

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	8,039	0.011 (0.037)	98.916
Belgium	11,333	0.017 (0.025)	200.816
Croatia	858	-0.015 (0.094)	13.153
Czech Republic	11,029	-0.018 (0.020)	342.321
Denmark	9,222	-0.016 (0.022)	279.523
England	44,647	-0.079 (0.020)***	765.176
Estonia	9,009	-0.044 (0.054)	43.081
France	10,725	-0.033 (0.019)*	408.578
Germany	9,061	-0.050 (0.021)**	262.482
Greece	5,403	-0.030 (0.092)	11.179
Hungary	616	-0.203 (0.122)*	36.201
Israel	3,782	-0.072 (0.076)	42.819
Italy	7,937	-0.102 (0.056)*	86.290
Luxembourg	1,624	-0.044 (0.051)	46.627
Netherlands	4,938	-0.046 (0.028)*	282.518
Poland	3,070	-0.133 (0.115)	34.864
Portugal	1,134	-0.617 (0.390)	3.455
Slovenia	5,438	0.015 (0.041)	68.322
Spain	6,613	-0.092 (0.037)**	162.698
Sweden	9,150	-0.003 (0.017)	221.178
Switzerland	7,015	0.015 (0.026)	159.139
Costa Rica	1,870	0.598 (0.574)	1.379
United States	75,515	-0.027 (0.036)	171.749
China	3,269	0.224 (0.157)	18.299
South Korea	21,527	0.154 (0.222)	7.252

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.17 Country-by-Country FEIV Models for the Effect of Retirement on Smoking

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	6,190	-0.101 (0.037)***	85.269
Belgium	8,196	-0.082 (0.025)***	138.372
Croatia	856	-0.085 (0.086)	13.146
Czech Republic	7,600	0.007 (0.020)	315.464
Denmark	6,313	-0.063 (0.027)**	230.762
England	44,416	0.008 (0.011)	760.243
Estonia	6,380	-0.048 (0.049)	29.340
France	8,548	0.005 (0.017)	324.711
Germany	5,623	0.024 (0.025)	163.204
Greece	4,295	-0.008 (0.132)	7.958
Hungary	616	0.206 (0.098)**	36.201
Israel	2,777	-0.028 (0.067)	32.812
Italy	5,750	-0.041 (0.040)	80.315
Luxembourg	728	-0.068 (0.079)	44.489
Netherlands	4,938	-0.012 (0.027)	282.518
Poland	2,115	-0.086 (0.103)	21.048
Slovenia	3,668	0.052 (0.037)	71.991
Spain	3,929	0.003 (0.039)	108.252
Sweden	6,325	-0.011 (0.025)	154.875
Switzerland	5,349	-0.030 (0.030)	95.367
Costa Rica	2,623	0.386 (0.247)	1.135
Mexico	21,848	0.100 (0.091)	15.535
United States	132,462	-0.009 (0.012)	395.455
China	7,379	-0.052 (0.053)	37.260
Japan	4,065	0.286 (0.296)	2.207
South Korea	21,526	-0.144 (0.122)	7.231

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.18 Country-by-Country FEIV Models for the Effect of Retirement on Binge Drinking

Country	Observations	Coef. (SE)	Kleibergen-Paap F
Austria	4,492	0.004 (0.040)	42.544
Belgium	4,804	-0.032 (0.051)	64.264
Czech Republic	5,706	0.065 (0.040)	137.946
Denmark	3,506	-0.002 (0.040)	144.161
England	33,054	-0.002 (0.022)	526.304
Estonia	4,740	0.032 (0.151)	8.853
France	5,250	0.042 (0.031)	100.022
Germany	1,899	-0.038 (0.041)	43.505
Israel	1,340	-0.028 (0.023)	31.520
Italy	3,194	0.019 (0.063)	28.817
Netherlands	2,908	-0.002 (0.050)	129.470
Poland	1,192	-0.123 (0.224)	8.670
Slovenia	1,998	0.141 (0.078)*	10.959
Spain	2,232	0.019 (0.054)	40.283
Sweden	2,798	-0.033 (0.035)	86.837
Switzerland	3,754	0.004 (0.038)	47.289
Mexico	21,627	0.126 (0.097)	15.175

United States	113,912	0.007 (0.009)	341.350
Japan	2,278	-0.347 (0.511)	0.805
South Korea	21,527	-0.022 (0.177)	7.252

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual and year. Robust standard errors clustering at individual and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.19 FEIV Models Excluding Countries with Weak IVs

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.051**	0.024***	0.143***	-0.038***	-0.009	0.005
	(0.020)	(0.006)	(0.018)	(0.009)	(0.006)	(0.009)
Observations	342,518	326,554	346,109	242,890	292,006	218,406
Kleibergen-Paap F	2411.781	2433.463	2375.399	2098.785	1775.578	901.990
Hansen J	0.053	0.279	0.855	3.806*	2.036	2.955*

Greece, Latvia, Malta, Portugal, Romania, Costa Rica, Japan, and South Korea are excluded from analysis. All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.20 FEIV Models Excluding Data from the United States

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.051**	0.019***	0.132***	-0.029***	-0.008	0.015
	(0.022)	(0.006)	(0.021)	(0.009)	(0.008)	(0.017)
Observations	251,993	257,382	251,522	197,309	192,057	128,299
Kleibergen-Paap F	2082.328	2120.431	2052.623	1844.518	1382.119	498.713
Hansen J	1.716	0.773	9.416***	0.221	4.273**	1.415

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.21 FEIV Models Excluding Countries Without Changing the SPA

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.052**	0.032***	0.149***	-0.030***	-0.004	0.013
	(0.022)	(0.006)	(0.020)	(0.011)	(0.007)	(0.010)
Observations	324,290	309,992	330,964	249,896	276,202	211,754
Kleibergen-Paap F	2028.669	2009.672	1979.637	1644.936	1486.158	763.661
Hansen J	0.558	0.060	2.041	3.547*	1.387	0.569

Cyprus, Finland, Luxembourg, Sweden, Switzerland, Costa Rica, Mexico, China, and Japan are excluded from analysis. All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.22 Stratified FEIV Models by Sex for the Effect of Retirement on the Raw Scores of Cognitive Function

	Men	Women
Retirement	-0.089	0.281***
	(0.101)	(0.087)
Observations	147,516	168,730
Kleibergen-Paap F	1027.651	1422.153
Hansen J	1.782	0.006

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.23 FEIV Models with Multiple Imputation

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.033	0.032***	0.141***	-0.025**	-0.012	0.014
	(0.022)	(0.006)	(0.020)	(0.010)	(0.009)	(0.010)
Observations	402164	402164	402164	317028	402164	253572
Kleibergen-Paap F	2282.936	2282.936	2282.936	2218.316	2282.936	844.731
Hansen J	0.530	0.142	3.451*	0.478	3.646*	0.982

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.24 Subgroup FEIV Models by Retirement Duration

	Cognitive function		Physical independence		Self-rated health	
	<5 years	≥5 years	<5 years	≥5 years	<5 years	≥5 years
Retirement	0.026	0.041	0.017**	0.046***	0.156***	0.121***
	(0.026)	(0.036)	(0.007)	(0.010)	(0.024)	(0.032)
Observations	277,378	300,186	264,810	288,898	284,966	307,455
Kleibergen-Paap F	1183.118	982.391	1151.145	1032.190	1169.870	960.784
Hansen J	0.072	0.158	1.233	0.779	3.395*	11.931***
	Physical inactivity		Smoking		Binge drinking	
	<5 years	≥5 years	<5 years	≥5 years	<5 years	≥5 years
Retirement	0.002	-0.041**	0.010	-0.001	0.026*	0.016
	(0.014)	(0.019)	(0.009)	(0.013)	(0.013)	(0.021)
Observations	190,124	208,611	239,323	256,064	180,880	191,665
Kleibergen-Paap F	741.853	740.301	775.172	534.881	437.062	247.862
Hansen J	0.190	4.007**	2.582	4.672**	0.009	0.271

Note: FEIV denotes fixed effect with instrumental variable. All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.25 First Stage Estimation for FEIV Models Including Interactions Between IVs and Marital Status

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Age	0.017***	0.017***	0.016***	0.021***	0.015***	0.013***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Age^2	0.003***	0.003***	0.003***	0.003***	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.001	-0.003	-0.003	-0.012***	-0.002	0.002
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
ERA	0.077***	0.079***	0.077***	0.076***	0.077***	0.063***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
ERA x Age	-0.006***	-0.007***	-0.006***	-0.008***	-0.008***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ERA x Age^2	-0.000***	-0.000***	-0.000***	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ERA x Married	0.011***	0.006	0.010***	0.014***	0.009**	0.002
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
ORA	0.174***	0.175***	0.168***	0.169***	0.159***	0.121***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
ORA x Age	-0.010***	-0.009***	-0.009***	-0.004***	-0.009***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
ORA x Age^2	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ORA x Married	0.008**	0.008**	0.008**	0.014***	0.008**	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Observations	377,276	362,973	384,631	272,824	324,519	242,211

Age squared was divided by 10 for ease of interpretation. All regressions are adjusted for fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table G.26 Second Stage Estimation for FEIV Models Including Interactions Between IVs and Marital Status

	Cognitive function	Physical independence	Self-rated health	Physical inactivity	Smoking	Binge drinking
Retirement	0.046**	0.027***	0.143***	-0.030***	-0.004	0.012
	(0.021)	(0.006)	(0.019)	(0.010)	(0.007)	(0.010)
Observations	377,276	362,973	384,631	272,824	324,519	242,211
Kleibergen-Paap F	1675.725	1668.328	1641.082	1395.281	1196.528	591.514
Hansen J	0.684	0.313	2.510	2.471	3.626*	0.640

All regressions are adjusted for age, age squared, marital status, and fixed effects of individual, country, year, and interactions between country and year. Robust standard errors clustering at individual, country, year, and interactions between country and year are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

H. Supplementary Materials for Chapter 3

Figure H.1 Sample Flowchart

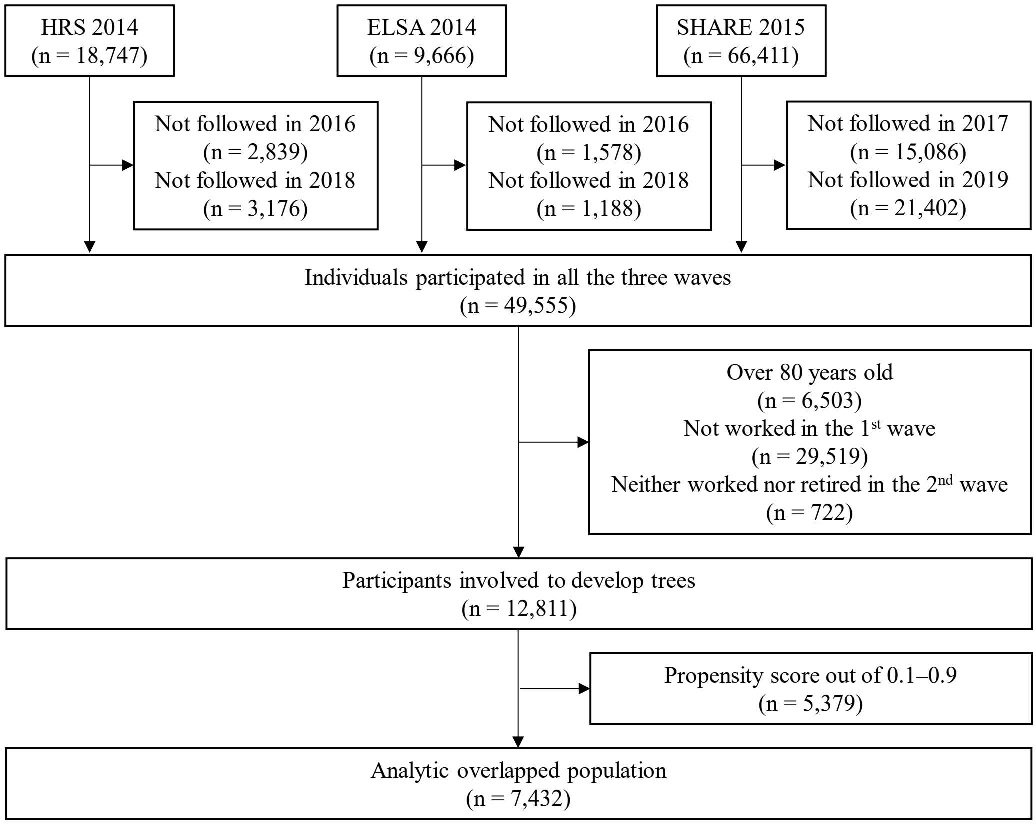


Table H.1 Occupational Codes

This study	HRS	ELSA	SHARE
	The 2010 Census	Standard Occupational Classification (2000)	1988 International Standard Classification of Occupations
Professional	1. Management occupations	1. Managers and senior officials	0. Armed forces
	2. Business and financial specialists	2. Managers and proprietors in agriculture and services	1. Legislator, senior official or manager
	3. Computer and mathematical occupations	3. Science and technology professionals	2. Professional
	4. Architecture and engineering occupations	4. Health professionals	3. Technician or associate professional
	5. Life, physical, and social science occupations	5. Teaching and research professionals	
	6. Community and social services occupations	6. Business and public service professionals	
	7. Legal occupations	7. Science and technology associate professionals	
	8. Education, training, and library occupations	8. Health and social welfare associate professionals	

	9. Arts, design, entertainment, sports, and media occupations	10. Culture, media and sports occupations	
	10. Healthcare practitioners and technical occupations	11. Business and public service associate professionals	
	11. Healthcare support occupations		
	23. Military specific occupations		
Clerk	17. Office and administrative support occupations	12. Administrative occupations	4. Clerk
		13. Secretarial and related occupations	
Service & sales	12. Protective service occupations	9. Protective service occupations	5. Service worker and shop and market sales worker
	13. Food preparation and serving occupations	18. Caring personal service occupations	
	14. Building and grounds cleaning and maintenance occupations	19. Leisure and other personal service occupations	
	15. Personal care and service occupations	20. Sales occupations	
	16. Sales occupations	21. Customer service occupations	
		25. Elementary administration and service occupations	
Manual labor	18. Farming, fishing, and forestry occupations	14. Skilled agricultural trades	6. Skilled agricultural or fishery worker
	19. Construction and extraction occupations	15. Skilled metal and electrical trades	7. Craft and related trades worker
	20. Installation, maintenance, and repair workers	16. Skilled construction and building trades	8. Plant and machine operator or assembler
	21. Production occupations	17. Textiles, printing and other skilled trades	9. Elementary occupation
	22. Transportation and material moving occupations	22. Process, plant and machine operatives	
		23. Transport and mobile machine drivers and operatives	
		24. Elementary trades, plant and storage related occupations	

Table H.2 Number of Imputed Values

Variable	Imputed values
Cognitive function	793
Retirement	0
Age	0
Men	0
Foreign-born	8
Education	210
Married	5
Living alone	0
No children	24
≥3 children	24
Asset	66
Income	955
Professional	1490
Clerk	1490
Service & sales	1490
Manual labor	1490
Physical demand	1715
Part-time job	223
Self-employed	13
Baseline cognition	268
Self-rated health	93
Depression	170
Life satisfaction	470
Hypertension	2
Diabetes	2
Cancer	2
Lung disease	2
Heart disease	2
Stroke	2
Arthritis	2
Psychiatric problems	2
Hyperlipemia	8
Health limitation in working	136
Difficulty in ADL	2
Difficulty in IADL	2
Distance eyesight	95
Near eyesight	99
Hearing	8
Pain problems	102
Obesity	844
Physical activity	14
Heavy drinking	331
Smoking	470

Figure H.2 Distribution of Cognitive Function

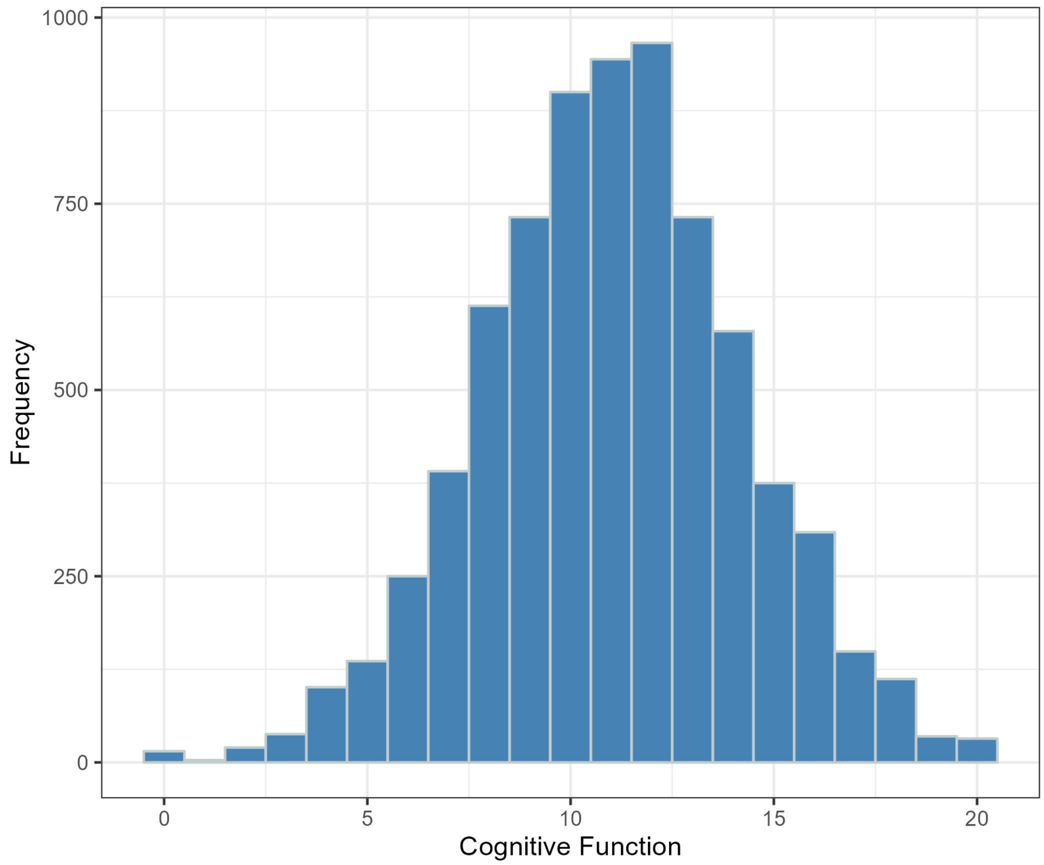


Figure H.3 Variable Importance

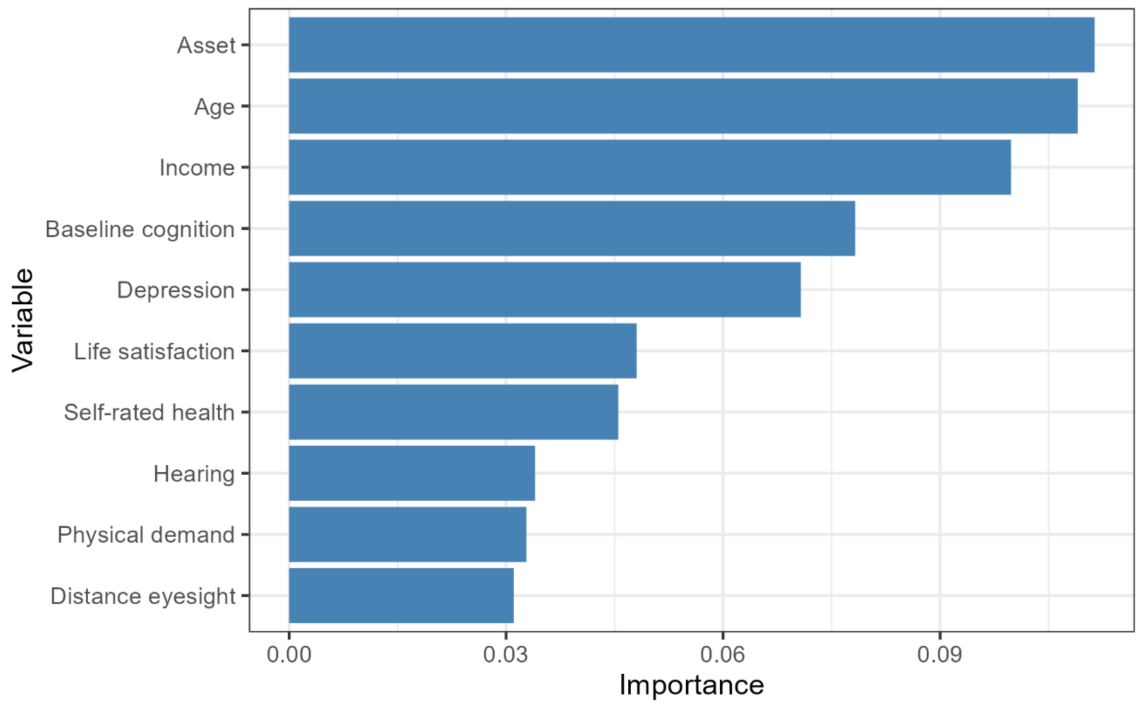


Figure H.4 Partial Dependence Plot for Continuous Variables

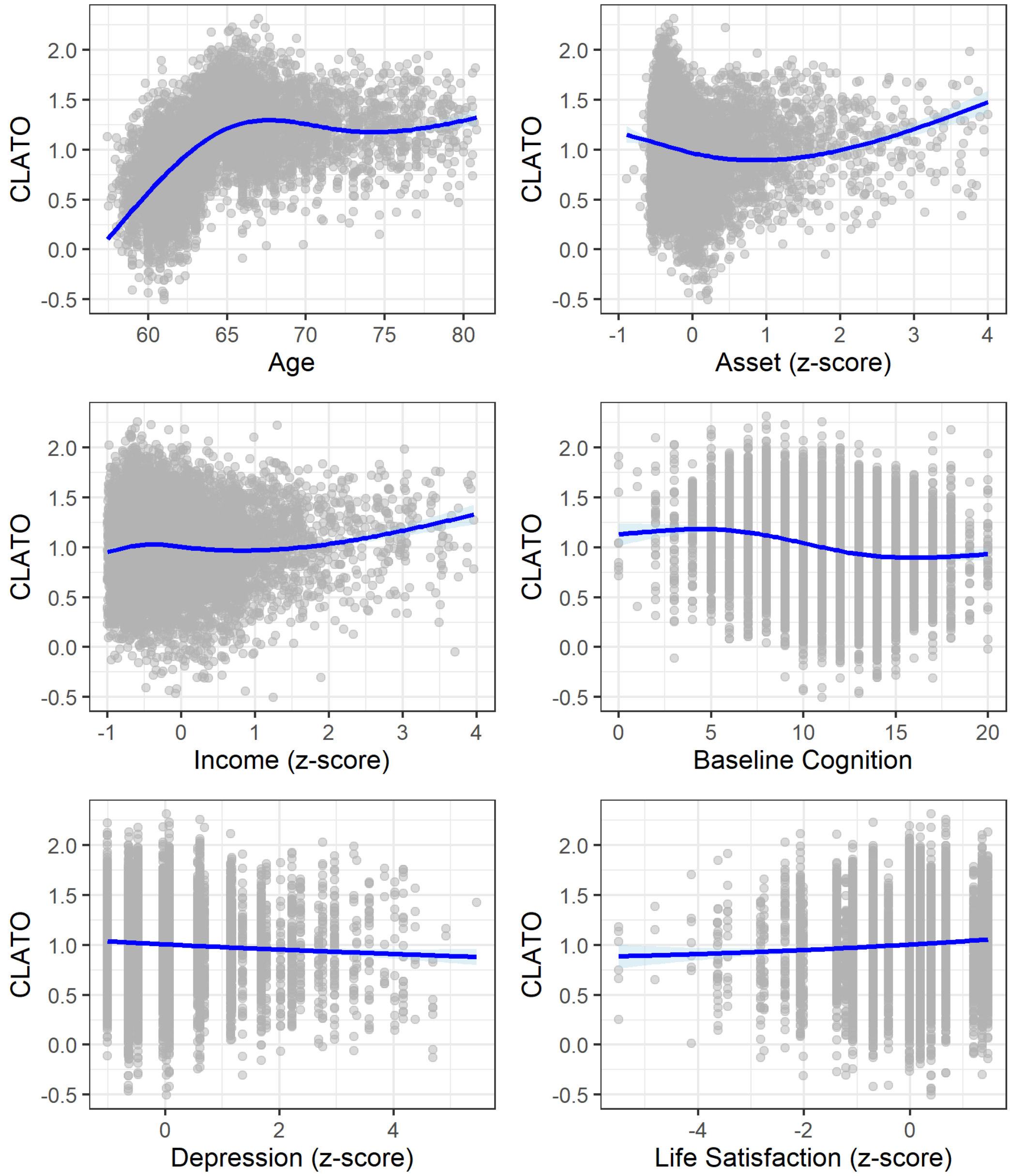


Table H.3 Monetary Cost Estimation in the United States

Ranking	Q1	Q2	Q3	Q4	Q5	Note
(1) Estimated CLATO	-0.579	0.767	1.163	1.379	1.662	From our estimates
(2) % of participants	20.04%	19.87%	19.96%	19.91%	20.22%	From our estimates
(3) Odds ratio for dementia risk	1.099	0.883	0.828	0.799	0.763	exp(ln(0.85) * (1)). Tierney et al. (2010) showed that a one-word increase in the RAVLT short delayed verbal recall was associated with 0.85 (95% CI: 0.78-0.92) times lower odds of dementia in five years.
(4) Population aged 55-59 in 2020			22,359,065			U.S. Census Bureau, the 2020 Census Demographic and Housing Characteristics File (DHC)
(5) % of workers at age 65			29.18%			RAND HRS in 2018
(6) Estimated # of workers at their age of 65			6,524,375			(4) * (5)
(7) Expected increase in # of workers at age 66			709,708			(6) / 9.2. Reaching the ORA increases the probability of retirement by 10.9% points in data from HRS. One out of 9.2 (= 1/0.109) workers would retire if they reached the ORA.
(8) Expected # of newly developed dementia	-14,036	16,525	24,392	28,373	33,968	(1 - (3)) * (2) * (7)
(9) Monetary cost of dementia per patient			$56,290			Hurd et al. (2013) estimated the monetary cost of dementia was $56,290 (95% CI: $42,746-$69,834) per person.
(10) Estimated additional cost	-$790M	$930M	$1,373M	$1,597M	$1,912M	(8) * (9)
(11) Total monetary cost			$5.0B (1.4%)			Hurd et al. (2013) estimated a total cost of $361B in 2030.

RAVLT stands for Rey Auditory Verbal Learning Test.

Table H.4 Monetary Cost Estimation in the United Kingdom

Ranking	Q1	Q2	Q3	Q4	Q5	Note
(1) Estimated CLATO	-0.579	0.767	1.163	1.379	1.662	From our estimates
(2) % of participants	20.48%	20.55%	20.12%	19.83%	19.03%	From our estimates
(3) Odds ratio for dementia risk	1.099	0.883	0.828	0.799	0.763	exp(ln(0.85) * (1)). Tierney et al. (2010) showed that a one-word increase in the RAVLT short delayed verbal recall was associated with 0.85 (95% CI: 0.78-0.92) times lower odds of dementia in five years.
(4) Population aged 55-59 in 2021			4,573,856			Office for National Statistics, Mid-Year Population Estimates June 2021
(5) % of workers at age 65			27.73%			ELSA in 2018
(6) Estimated # of workers at their age of 65			1,268,330			(4) * (5)
(7) Expected increase in # of workers at age 66			552,727			(6) / 2.3. Reaching the ORA increases the probability of retirement by 43.6% points in data from ELSA. One out of 2.3 (= 1/0.436) workers would retire if they reached the ORA.
(8) Expected # of newly developed dementia	-11,169	13,313	19,149	22,001	24,893	(1 - (3)) * (2) * (7)
(9) Monetary cost of dementia per patient			£47,997			£59,200M / 1,233,400. Wittenberg et al. (2019) estimated the number of people with dementia and its total cost in 2030.
(10) Estimated additional cost	-£536M	£639M	£919M	£1,056M	£1,195M	(8) * (9)
(11) Total monetary cost			£3.3B (5.2%)			Wittenberg et al. (2019) estimated a total cost of £59.2B in 2030.

RAVLT stands for Rey Auditory Verbal Learning Test.