WINPEC Working Paper Series No. E2306
September 2023

Heterogeneous Treatment Effect of Retirement on Cognitive Function

SATO Koryu, NOGUCHI Haruko, and INOUE Kosuke

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

Heterogeneous Treatment Effect of Retirement on Cognitive Function

By KORYU SATO, HARUKO NOGUCHI, AND KOSUKE INOUE*

This study used instrumental variable causal forests to explore the heterogeneous treatment effect of retirement on cognitive function using data from 19 countries. We found that, on average, retirees have better cognitive function than workers and that the conditional average treatment effects vary depending on individuals' characteristics. Policymakers should provide early retirement options in the pension system to allow individuals to decide when to retire. The balance between the social benefits of raising the state pension age and the individual costs of increasing the risk of dementia by delaying retirement should be considered. (*JEL* I10, J26, C26)

* Sato: Department of Social Epidemiology, Graduate School of Medicine and School of Public Health, Kyoto University, Room 204, Science Frontier Laboratory, Yoshida-Konoecho, Sakyo-ku, Kyoto-shi, Kyoto 606-8315, Japan; Graduate School of Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan (email: sato.koryu.8i@kyoto-u.ac.jp). Noguchi: Graduate School of Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan (email: h.noguchi@waseda.jp). Inoue: the Hakubi Project, Department of Social Epidemiology, Graduate School of Medicine and School of Public Health, Kyoto University, Room 204, Science Frontier Laboratory, Yoshida-Konoecho, Sakyo-ku, Kyoto-shi, Kyoto 606-8315, Japan (email: inoue.kosuke.2j@kyoto-u.ac.jp).
This study was 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 data; 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 analysis 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; RAND HRS Longitudinal File 2018 (V2); and the Harmonized HRS dataset and Codebook, Version C as of January 2022. 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 harmonized datasets and more information are available through the Gateway to Global Aging Data website (https://g2aging.org/). This study 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. 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 (Hurd et al. 2013; Wimo et al. 2017; Cimler et al. 2019; 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 and 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.[1] Retirement dramatically changes individuals' budget 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

[1] 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).

attributing it to issues with model specifications (Nishimura, Oikawa, and Motegi 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, Oikawa, and Motegi 2018; Rose 2020), whereas analysis using data from the Korean Longitudinal Study of Aging indicated a beneficial effect (Nishimura, Oikawa, and Motegi 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 II describes the data used in this study, Section III presents the empirical model, Section IV reports the results, and Section V discusses the results and concludes the paper.

II. Data

A. 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).[2] 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.[3]

Appendix A 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

[2] 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 and Noguchi 2023).

[3] 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.

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 for the development of IV forests.

B. 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 (Bonsang, Adam, and Perelman 2012; Coe and Zamarro 2011; Coe et al. 2012; Bianchini and Borella 2016; Celidoni, Dal Bianco, and Weber 2017; Nishimura, Oikawa, and Motegi 2018; Rose 2020).

C. Retirement and State Pension Age

Labor force status was self-reported in the surveys, as described in Appendix B 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 (Bianchini and Borella 2016; Atalay, Barrett, and Staneva 2019). Other studies have defined retirement as not working (Coe and Zamarro 2011; Bonsang, Adam, and Perelman 2012; Bingley and Martinello 2013), 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 and 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 C 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 1 (men) and 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.

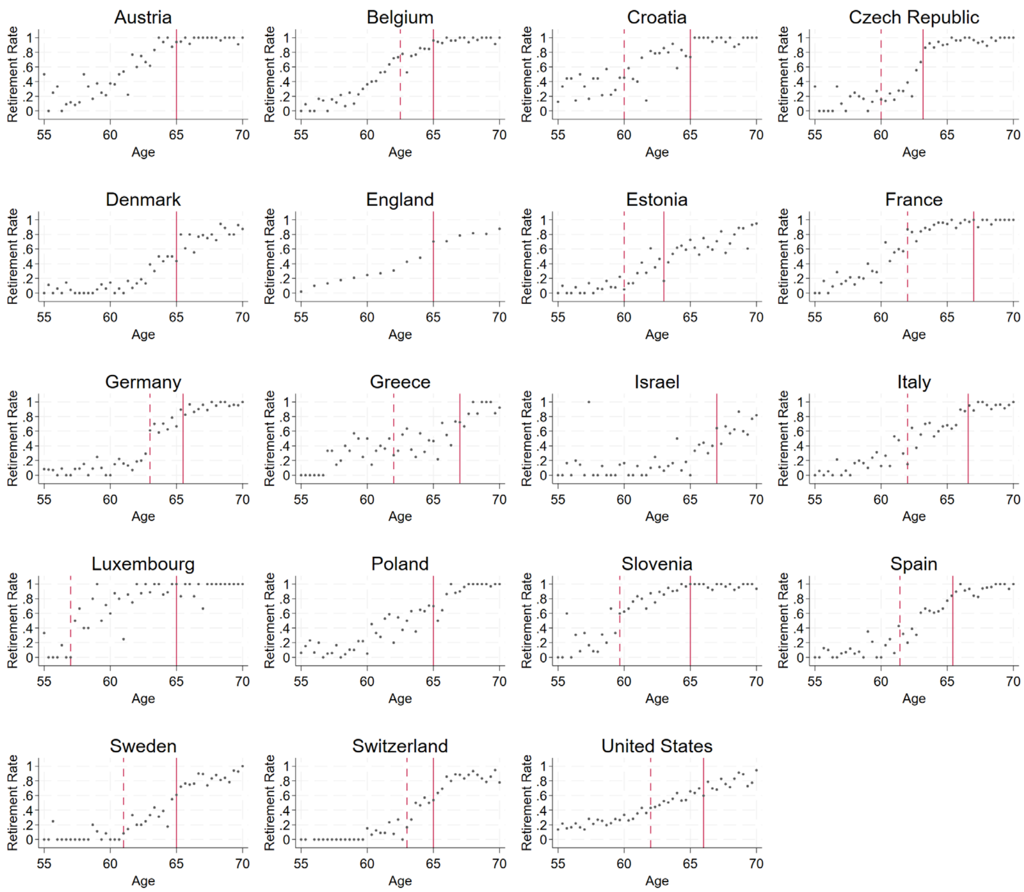

FIGURE 1. RETIREMENT RATE OF MEN

Notes: 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.

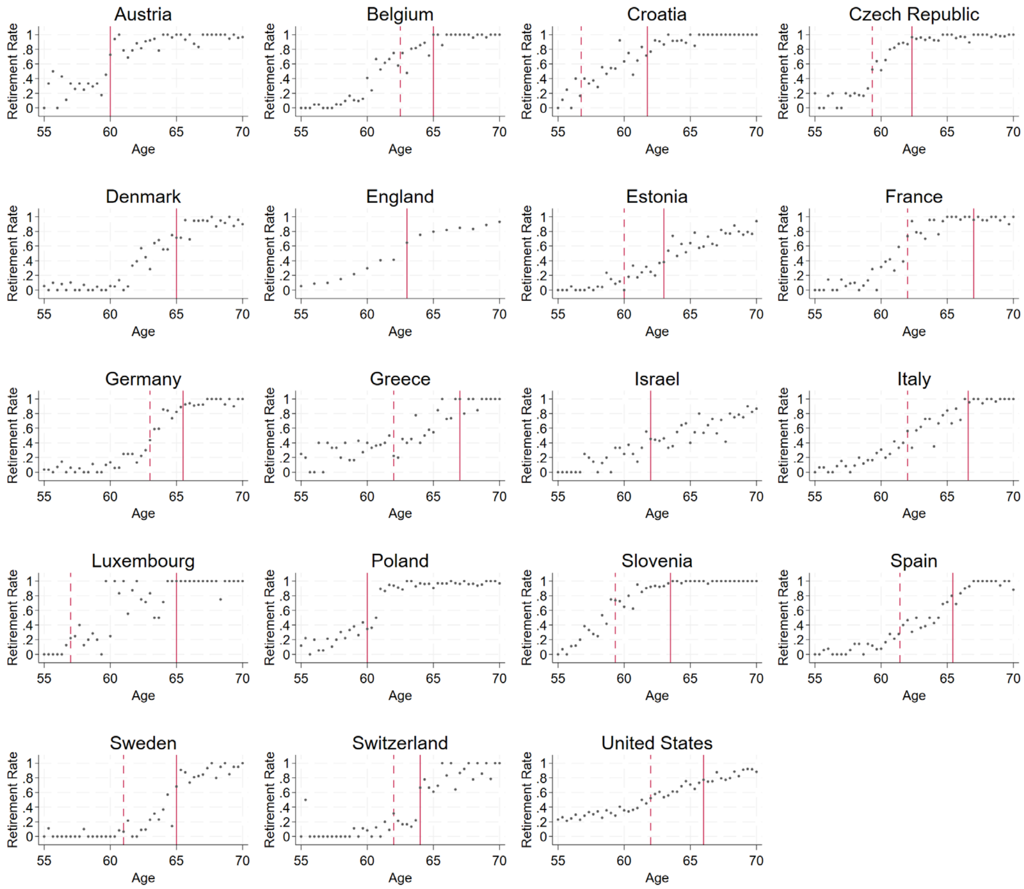

FIGURE 2. RETIREMENT RATE OF WOMEN

Notes: 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.

D. Covariates

We considered 60 harmonized covariates obtained from the first wave to develop IV forests. Table 1 presents definitions of these covariates.

TABLE 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 and 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 and Lee (2012a) for details about the harmonization of income measures.
Occupation	Binary	We created four 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 D 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 being self-employed and 0 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 participants' 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 and 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 being troubled with pain and 0 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 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.

Notes: HRS stands for the Health and Retirement Study, ELSA stands for the English Longitudinal Study on Ageing, and SHARE stands for the Survey of Health, Ageing and Retirement in Europe. ADL and IADL stand for activities of daily living and instrumental activities of daily living, respectively.

III. Empirical Model

A. 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

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

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

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

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 $b = 1, \dots, B$

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

$L_b(x)$ 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 and Imbens 2016; Wager and Athey 2018; Athey, Tibshirani, and Wager 2019). 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

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

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

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

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

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

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

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

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 $\alpha_i(x)$.

B. Statistical Analysis

Our IV forests developed 2000 trees and tuned their parameters[4] 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 i th observation) of these marginal expectations, $\hat{y}^{(-i)}, \widehat{w}^{(-i)}, \hat{z}^{(-i)}$, 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 $\widehat{w}^{(-i)}$ 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 $\tilde{Y}_i = Y_i - \hat{y}^{(-i)}(X_i)$, $\widetilde{W}_i = W_i - \widehat{w}^{(-i)}(X_i)$, 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 $\{\tilde{Y}_i, \widetilde{W}_i, \tilde{Z}_i\}_{i=1}^{n} \in \mathcal{P}$.[5] 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.

[4] 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.

[5] 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.

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 the 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; subgroup of individuals who received the least benefits from retirement) to Q5 (the highest CLATO; subgroup of individuals who received the most benefits from retirement). This procedure was repeated for each iteration.[6] 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.[7] 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, Moineddin, and McDowell 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)

[6] Namely, we performed a 10-fold cross-fitting procedure.

[7] 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.

provided monetary cost estimates for dementia care per patient and their projections for 2030 in the 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)[8], assuming that the data were missing at random. Appendix E reports the imputed values for each variable.

IV. Results

A. 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 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

[8] 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.

composition of countries. The outcomes appeared to be normally distributed, as shown in Appendix F.

TABLE 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
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

Notes: ADL and IADL stand for activities of daily living and instrumental activities of daily living, respectively. Imputed data are used.

B. Average Treatment Effects

Table 3 compares the LATO estimated using IV forests with the estimates of the conventional OLS, 2SLS, and non-IV forests. 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—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
F statistic		163.037***		
Sargan statistic		1.177		

Notes: OLS stands for ordinary least squares, 2SLS stands for two-stage least squares, IV stands for instrumental variable, ERA stands for early retirement age, and ORA stands for official retirement age. Standard errors are shown in parentheses.

[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.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

C. Heterogeneous Treatment Effect of Retirement

Figure 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.

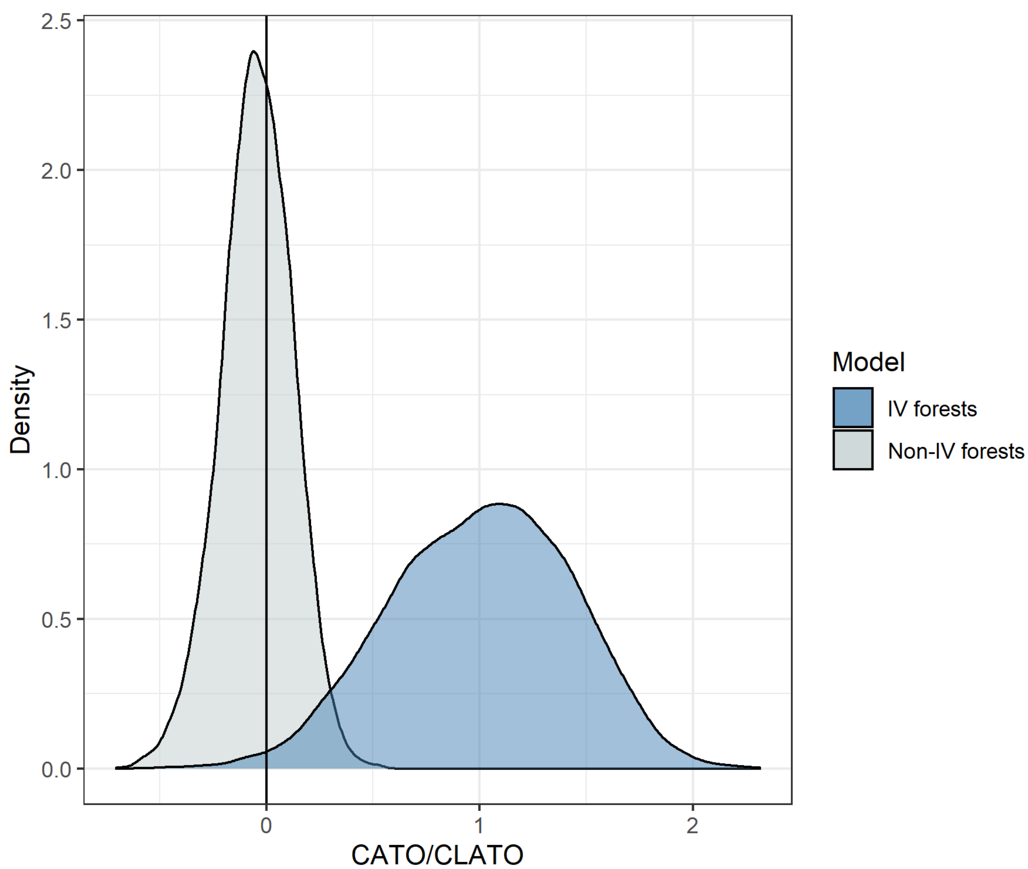

FIGURE 3. DISTRIBUTION OF CONDITIONAL AVERAGE TREATMENT EFFECTS

Notes: IV stands for an instrumental variable. 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 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.

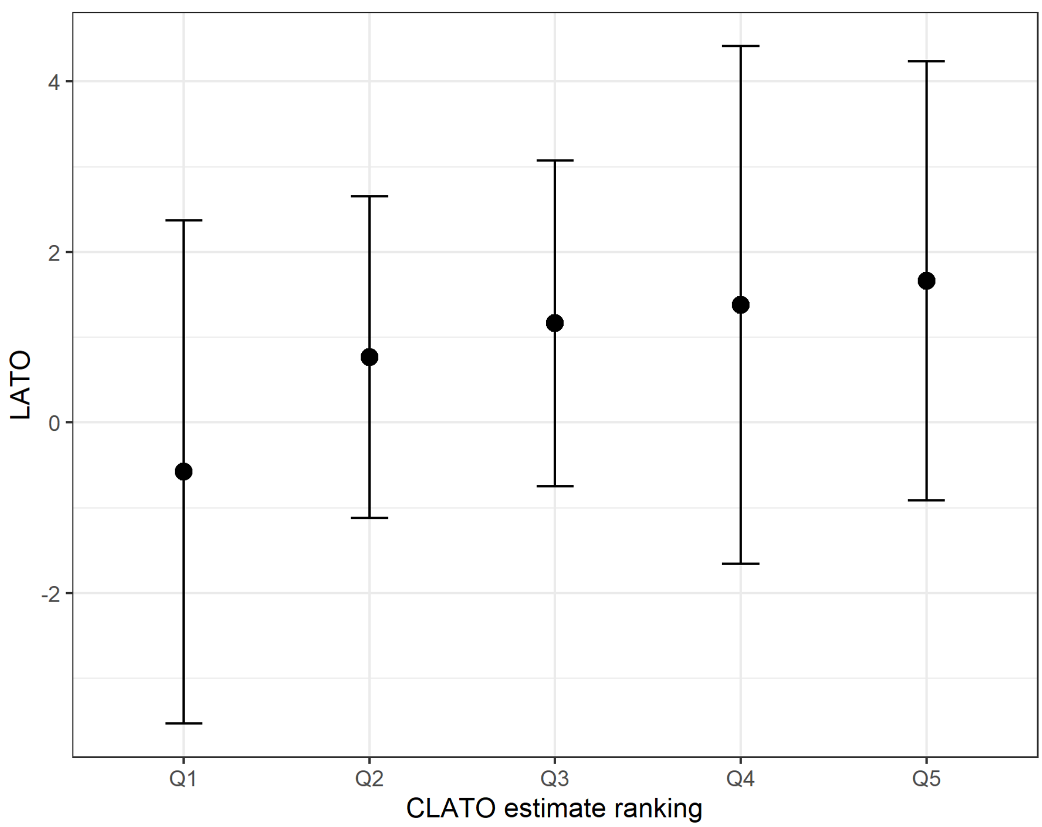

FIGURE 4. CALIBRATION PLOT FOR CLATO

Notes: CLATO stands for the conditional local average treatment effect on the overlap population.

Figure 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 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 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.

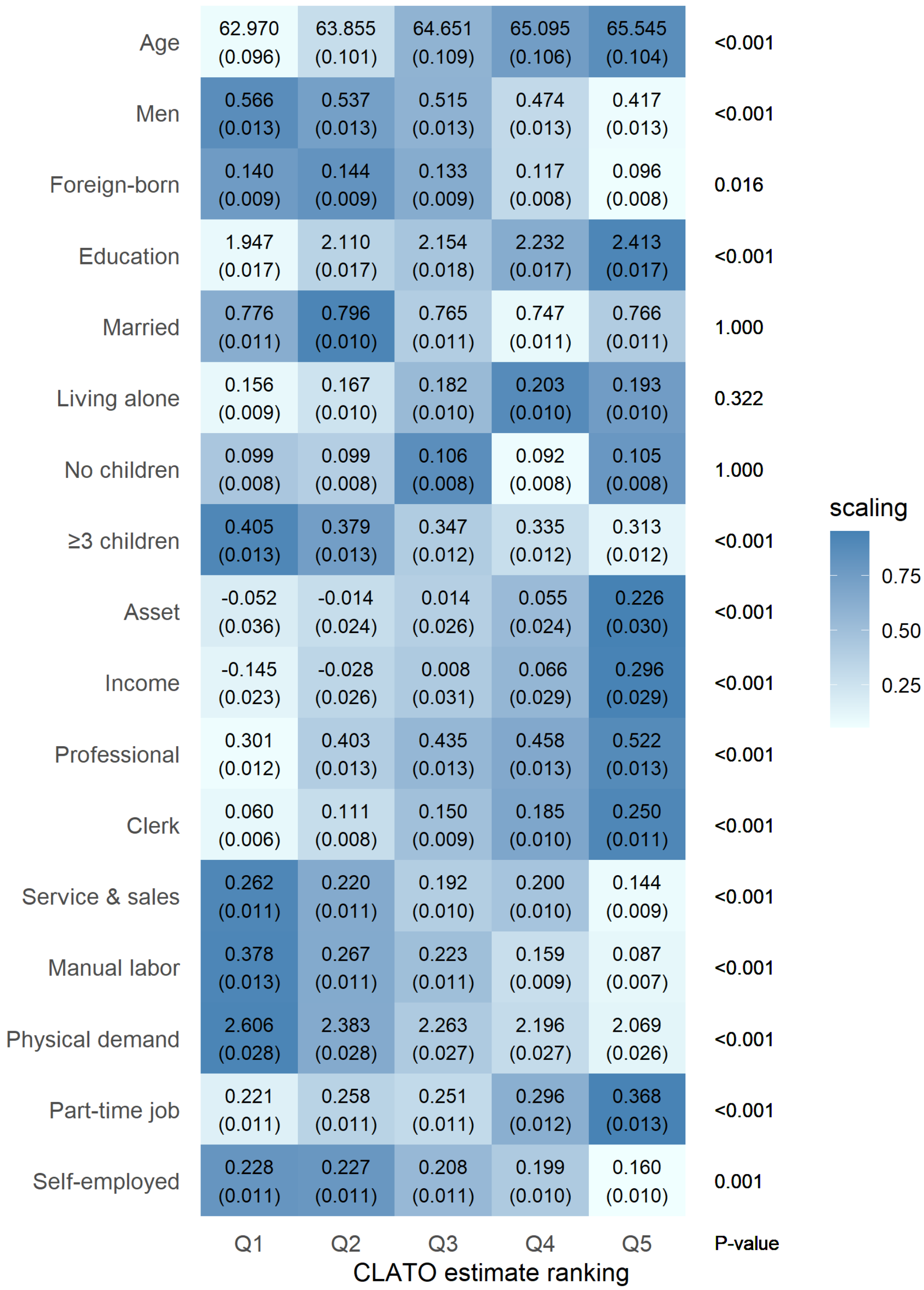

FIGURE 5. HETEROGENEITY IN INDIVIDUAL CHARACTERISTICS

Notes: CLATO stands for the conditional local average treatment effect on the overlap population. 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.

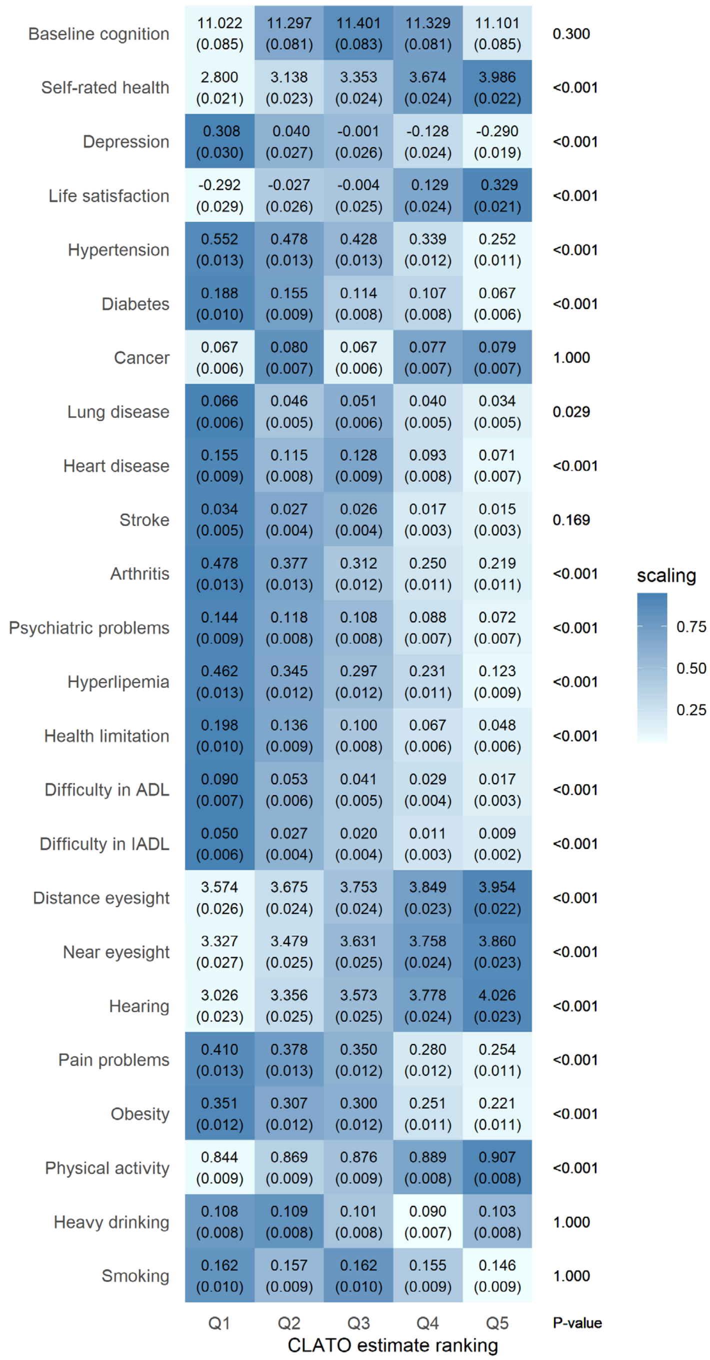

FIGURE 6. HETEROGENEITY IN HEALTH AND BEHAVIORS

Notes: ADL and IADL stand for activities of daily living and instrumental activities of daily living, respectively. 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.

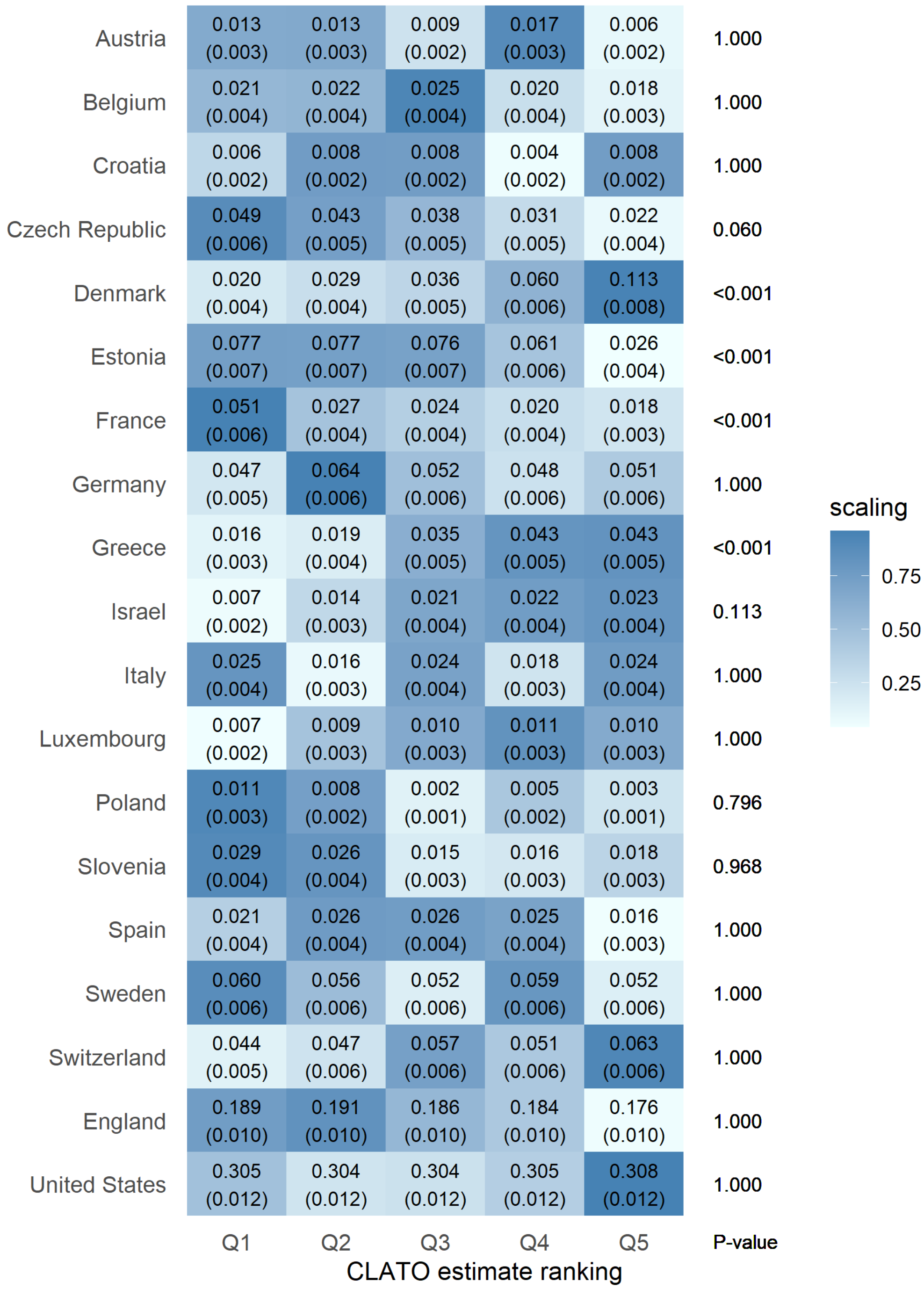

FIGURE 7. HETEROGENEITY IN COUNTRIES

Notes: 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.

Appendix G 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.

D. Monetary Cost Estimation of Increasing ORA

Based on the LATO estimates for the quintile groups shown in Figure 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 H 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.

E. Robustness Checks

Table 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 4, all the LATO estimates are similar to our main results, which suggests that our findings are robust, even in different settings.

TABLE 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

Notes: 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.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

V. 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 (de Grip et al. 2015; Bianchini and Borella 2016). 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.[9] Improvements in cognitive function after retirement can be explained through several pathways. First, job strain is a potential risk factor for decreased cognitive function (Elovainio et al. 2009; Agbenyikey et al. 2015), 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 (Kämpfen and Maurer 2016; Myllyntausta et al. 2018; Müller and Shaikh 2018; Celidoni and Rebba 2017; Kesavayuth, Rosenman, and Zikos 2018; Eibich 2015; Syse et al. 2017; Insler 2014), 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 (Sato and Noguchi 2023; Atalay, Barrett, and Staneva 2019). 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.

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

[9] 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.

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 (Fisher et al. 2014; Andel et al. 2015; Kajitani, Sakata, and Mckenzie 2017; Romero Starke et al. 2019; Carr et al. 2020; Vélez-Coto et al. 2021). These findings suggest that a mentally demanding job has a protective effect on cognitive function, which 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, Moineddin, and McDowell 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 and Lee 2012a; 2012b; Hu and Lee 2012; Shih, Lee, and Das 2012; Zamarro and 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. 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

Agbenyikey, W., R. Karasek, M. Cifuentes, P. A. Wolf, S. Seshadri, J. A. Taylor, A. S. Beiser, and R. Au. 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, Ross, Frank J. Infurna, Elizabeth A. Hahn Rickenbach, Michael Crowe, Lisa Marchiondo, and Gwenith G. Fisher. 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): 442–46. https://doi.org/10.1136/jech-2014-204754.

Angrisani, Marco, and Jinkook Lee. 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.

———. 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., and I. Hertz-Picciotto. 1994. "The Evolving Concept of the Healthy Worker Survivor Effect." *Epidemiology* 5 (2): 189–96. https://doi.org/10.1097/00001648-199403000-00009.

Atalay, Kadir, Garry F. Barrett, and Anita Staneva. 2019. "The Effect of Retirement on Elderly Cognitive Functioning." *Journal of Health Economics* 66 (July): 37–53. https://doi.org/10.1016/j.jhealeco.2019.04.006.

Athey, Susan, and Guido Imbens. 2016. "Recursive Partitioning for Heterogeneous Causal Effects." *Proceedings of the National Academy of Sciences of the United States of America* 113 (27): 7353–60. https://doi.org/10.1073/pnas.1510489113.

Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. "Generalized Random Forests." *The Annals of Statistics* 47 (2): 1148–78. https://doi.org/10.1214/18-AOS1709.

Bianchini, Laura, and Margherita Borella. 2016. "Retirement and Memory in Europe." *Ageing & Society* 36: 1434–58. https://doi.org/10.1017/S0144686X15000434.

Biewen, Martin, and Philipp Kugler. 2021. "Two-Stage Least Squares Random Forests with an Application to Angrist and Evans (1998)." *Economics Letters* 204 (July): 109893. https://doi.org/10.1016/j.econlet.2021.109893.

Bingley, Paul, and Alessandro Martinello. 2013. "Mental Retirement and Schooling." *European Economic Review* 63 (October): 292–98. https://doi.org/10.1016/j.euroecorev.2013.01.004.

Bonsang, Eric, Stéphane Adam, and Sergio Perelman. 2012. "Does Retirement Affect Cognitive Functioning?" *Journal of Health Economics* 31 (3): 490–501. https://doi.org/10.1016/j.jhealeco.2012.03.005.

Carr, Dawn C., Robert Willis, Ben Lennox Kail, and Laura L. Carstensen. 2020. "Alternative Retirement Paths and Cognitive Performance: Exploring the Role of Preretirement Job Complexity." *The Gerontologist* 60 (3): 460–71. https://doi.org/10.1093/geront/gnz079.

Celidoni, Martina, Chiara Dal Bianco, and Guglielmo Weber. 2017. “Retirement and Cognitive Decline. A Longitudinal Analysis Using SHARE Data.” *Journal of Health Economics* 56 (December): 113–25. https://doi.org/10.1016/j.jhealeco.2017.09.003.

Celidoni, Martina, and Vincenzo Rebba. 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): 805–30. https://doi.org/10.1007/s10198-016-0828-8.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. “Double/Debiased Machine Learning for Treatment and Structural Parameters.” *The Econometrics Journal* 21 (1): C1–68. https://doi.org/10.1111/ectj.12097.

Cimler, Richard, Petra Maresova, Jitka Kuhnova, and Kamil Kuca. 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, Norma B., Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen Maurer. 2012. “The Effect of Retirement on Cognitive Functioning.” *Health Economics* 21 (8): 913–27. https://doi.org/10.1002/hec.1771.

Coe, Norma B., and Gema Zamarro. 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.

Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik. 2009. “Dealing with Limited Overlap in Estimation of Average Treatment Effects.” *Biometrika* 96 (1): 187–99. https://doi.org/10.1093/biomet/asn055.

Denier, Nicole, Sean A. P. Clouston, Marcus Richards, and Scott M. Hofer. 2017. “Retirement and Cognition: A Life Course View.” *Advances in Life Course Research* 31 (March): 11–21. https://doi.org/10.1016/j.alcr.2016.10.004.

Eibich, Peter. 2015. “Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity.” *Journal of Health Economics* 43 (September): 1–12. https://doi.org/10.1016/j.jhealeco.2015.05.001.

Elovainio, M., J. E. Ferrie, A. Singh-Manoux, D. Gimeno, R. De Vogli, M. J. Shipley, J. Vahtera, E. J. Brunner, M. G. Marmot, and M. Kivimäki. 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, Gwenith G., Alicia Stachowski, Frank J. Infurna, Jessica D. Faul, James Grosch, and Lois E. Tetrick. 2014. “Mental Work Demands, Retirement, and Longitudinal Trajectories of Cognitive Functioning.” *Journal of*

Occupational Health Psychology 19 (2): 231–42. https://doi.org/10.1037/a0035724.
Grip, Andries de, Arnaud Dupuy, Jelle Jolles, and Martin van Boxtel. 2015. "Retirement and Cognitive Development in the Netherlands: Are the Retired Really Inactive?" *Economics and Human Biology* 19 (December): 157–69. https://doi.org/10.1016/j.ehb.2015.08.004.
Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2): 223–55.
Hu, Peifeng, and Jinkook Lee. 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, Michael D., Paco Martorell, Adeline Delavande, Kathleen J. Mullen, and Kenneth M. Langa. 2013. "Monetary Costs of Dementia in the United States." *The New England Journal of Medicine* 368 (14): 1326–34. https://doi.org/10.1056/NEJMsa1204629.
Insler, Michael. 2014. "The Health Consequences of Retirement." *Journal of Human Resources* 49 (1): 195–233. https://doi.org/10.3368/jhr.49.1.195.
Kajitani, Shinya, Kei Sakata, and Colin Mckenzie. 2017. "Occupation, Retirement and Cognitive Functioning." *Ageing & Society* 37 (8): 1568–96. https://doi.org/10.1017/S0144686X16000465.
Kämpfen, Fabrice, and Jürgen Maurer. 2016. "Time to Burn (Calories)? The Impact of Retirement on Physical Activity among Mature Americans." *Journal of Health Economics* 45 (January): 91–102. https://doi.org/10.1016/j.jhealeco.2015.12.001.
Kesavayuth, Dusanee, Robert E. Rosenman, and Vasileios Zikos. 2018. "Retirement and Health Behaviour." *Applied Economics* 50 (54): 5859–76. https://doi.org/10.1080/00036846.2018.1488070.
Kraft, Matthew A. 2020. "Interpreting Effect Sizes of Education Interventions." *Educational Researcher* 49 (4): 241–53. https://doi.org/10.3102/0013189X20912798.
Lee, Jinkook, Drystan Phillips, Jenny Wilkens, and 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): S5–16. https://doi.org/10.1093/geronb/gbab050.
Li, Zhe. 2022. "The Social Security Retirement Age." CRC Report R44670. Washington, DC: Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R44670.
Livingston, Gill, Jonathan Huntley, Andrew Sommerlad, David Ames, Clive Ballard, Sube Banerjee, Carol Brayne, et al. 2020. "Dementia Prevention,

Intervention, and Care: 2020 Report of the Lancet Commission." *The Lancet* 396 (10248): 413–46. https://doi.org/10.1016/S0140-6736(20)30367-6.

Mayer, Michael. 2021. "MissRanger: Fast Imputation of Missing Values." https://cran.r-project.org/web/packages/missRanger/missRanger.pdf.

Mazzonna, Fabrizio, and Franco Peracchi. 2012. "Ageing, Cognitive Abilities and Retirement." *European Economic Review* 56 (4): 691–710. https://doi.org/10.1016/j.euroecorev.2012.03.004.

Morris, J. C., A. Heyman, R. C. Mohs, J. P. Hughes, G. van Belle, G. Fillenbaum, E. D. Mellits, and C. Clark. 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): 1159–65. https://doi.org/10.1212/wnl.39.9.1159.

Müller, Tobias, and Mujaheed Shaikh. 2018. "Your Retirement and My Health Behavior: Evidence on Retirement Externalities from a Fuzzy Regression Discontinuity Design." *Journal of Health Economics* 57 (January): 45–59. https://doi.org/10.1016/j.jhealeco.2017.10.005.

Myllyntausta, Saana, Paula Salo, Erkki Kronholm, Jaana Pentti, Mika Kivimäki, Jussi Vahtera, and Sari Stenholm. 2018. "Changes in Sleep Difficulties During the Transition to Statutory Retirement." *Sleep* 41 (1): zsx182. https://doi.org/10.1093/sleep/zsx182.

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

Nishimura, Yoshinori, Masato Oikawa, and Hiroyuki Motegi. 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.

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

Robinson, Peter. 1988. "Root- N-Consistent Semiparametric Regression." *Econometrica* 56 (4): 931–54.

Romero Starke, Karla, Andreas Seidler, Janice Hegewald, Anna Klimova, and Katie Palmer. 2019. "Retirement and Decline in Episodic Memory: Analysis from a Prospective Study of Adults in England." *International Journal of Epidemiology* 48 (6): 1925–36. https://doi.org/10.1093/ije/dyz135.

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

Sato, Koryu, and Haruko Noguchi. 2023. “Impact of Retirement on Health: Evidence from 35 Countries.” WINPEC Working Paper Series No. E2301. Tokyo, Japan. https://doi.org/10.2139/ssrn.4413645.

Shih, Regina A., Jinkook Lee, and Lopamudra Das. 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.

Syse, Astri, Marijke Veenstra, Trude Furunes, Reidar J. Mykletun, and Per Erik Solem. 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.

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

Tierney, Mary C., Rahim Moineddin, and Ian McDowell. 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): 1231–40. https://doi.org/10.3233/JAD-2010-100516.

Tulving, Endel. 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.

United Kingdom Government. 2014. “State Pension Age Timetable.” GOV.UK. May 15, 2014. https://www.gov.uk/government/publications/state-pension-age-timetable.

Vélez-Coto, María, Ross Andel, Miguel Pérez-García, and Alfonso Caracuel. 2021. “Complexity of Work with People: Associations with Cognitive Functioning and Change after Retirement.” *Psychology and Aging* 36: 143–57. https://doi.org/10.1037/pag0000584.

Wager, Stefan, and Susan Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” *Journal of the American Statistical Association* 113 (523): 1228–42. https://doi.org/10.1080/01621459.2017.1319839.

Wimo, Anders, Maëlenn Guerchet, Gemma-Claire Ali, Yu-Tzu Wu, A. Matthew Prina, Bengt Winblad, Linus Jönsson, Zhaorui Liu, and Martin Prince. 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, Raphael, Bo Hu, Luis Barraza-Araiza, and Amritpal Rehill. 2019. "Projections of Older People with Dementia and Costs of Dementia Care in the United Kingdom, 2019–2040." CPEC Working Paper 5. London: 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. 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/.

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

Yadlowsky, Steve, Scott Fleming, Nigam Shah, Emma Brunskill, and Stefan Wager. 2021. "Evaluating Treatment Prioritization Rules via Rank-Weighted Average Treatment Effects." arXiv. https://doi.org/10.48550/arXiv.2111.07966.

Zamarro, Gema, and Jinkook Lee. 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.

Appendixes

A. Sample Flowchart

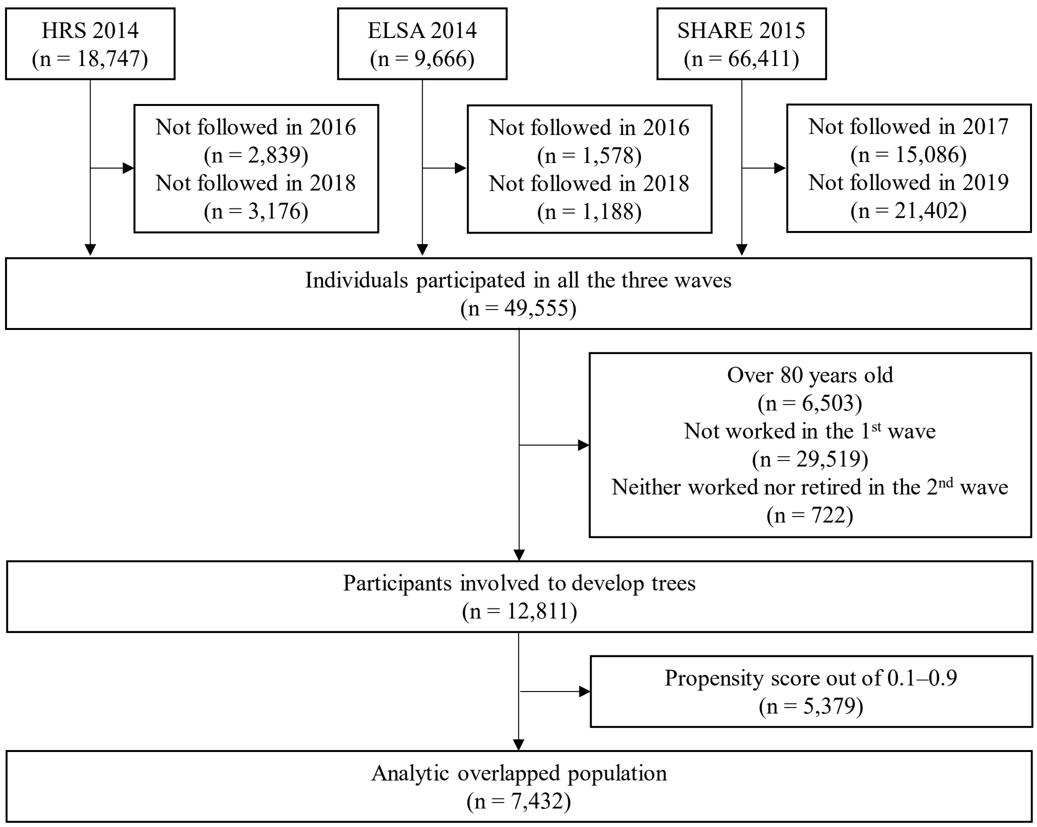

FIGURE A. SAMPLE FLOWCHART

Notes: HRS stands for the Health and Retirement Study, ELSA stands for the English Longitudinal Study on Ageing, and SHARE stands for the Survey of Health, Ageing and Retirement in Europe.

B. Measurement of Labor Force Status

The Health and Retirement Study (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".

The English Longitudinal Study on Ageing (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.

The Survey of Health, Ageing and Retirement in Europe (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.

TABLE A—SUMMARY OF THE HARMONIZED VARIABLE OF LABOR FORCE STATUS

This study	HRS	ELSA	SHARE
Included as workers	1. working full-time	1. employed	1. employed or self-employed
	2. working part-time	2. self-employed	
Included as retirees	4. partly retired	4. partly retired	5. retired
	5. retired	5. retired	
Excluded from analyses	3. unemployed	3. unemployed	3. unemployed
	6. disabled	6. disabled	6. permanently sick or disabled
	7. not in the labor force	7. looking after home or family	8. homemaker

Notes: HRS stands for the Health and Retirement Study, ELSA stands for the English Longitudinal Study on Ageing, and SHARE stands for the Survey of Health, Ageing and Retirement in Europe. Retirement status was determined based on the variable of labor force status (RwLBRF) in the harmonized datasets.

C. Early and Official Retirement Age

TABLE B—EARLY AND OFFICIAL RETIREMENT AGE

Country	Year	Men		Women	
		ERA	ORA	ERA	ORA
Austria	2017	NA	65	NA	60
Belgium	2017	62.5	65	62.5	65
Croatia	2017	60	65	56.75	61.75
Czech Republic[a]	2017	60	63.17	59.33	62.33
Denmark	2017	NA	65	NA	65
England	2016	NA	65	NA	63
Estonia	2017	60	63	60	63
France	2017	62	67	62	67
Germany	2017	63	65.5	63	65.5
Greece	2017	62	67	62	67
Israel	2017	NA	67	NA	62
Italy	2017	62	66.58	62	66.58

Luxembourg	2017	57	65	57	65
Poland	2017	NA	65	NA	60
Slovenia	2017	59.67	65	59.33	63.5
Spain	2017	61.42	65.42	61.42	65.42
Sweden	2017	61	65	61	65
Switzerland	2017	63	65	62	64
United States	2016	62	66	62	66

Notes: ERA and ORA stand for early and official retirement age, respectively. NA denotes not applicable.

Source: The United States Social Security Administration "Social Security Programs Throughout the World"; OECD "Pensions at a Glance"; websites of the authorities of each country.

[a] ORA for women is determined according to the number of raised children. ORA of 62 years and 4 months is for women without children in 2017.

D. Occupational Codes

TABLE C—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	

Notes: HRS stands for the Health and Retirement Study, ELSA stands for the English Longitudinal Study on Ageing, and SHARE stands for the Survey of Health, Ageing and Retirement in Europe.

E. Number of Imputed Values

TABLE D—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 limitations 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

Notes: ADL and IADL stand for activities of daily living and instrumental activities of daily living, respectively.

F. Distribution of Cognitive Function

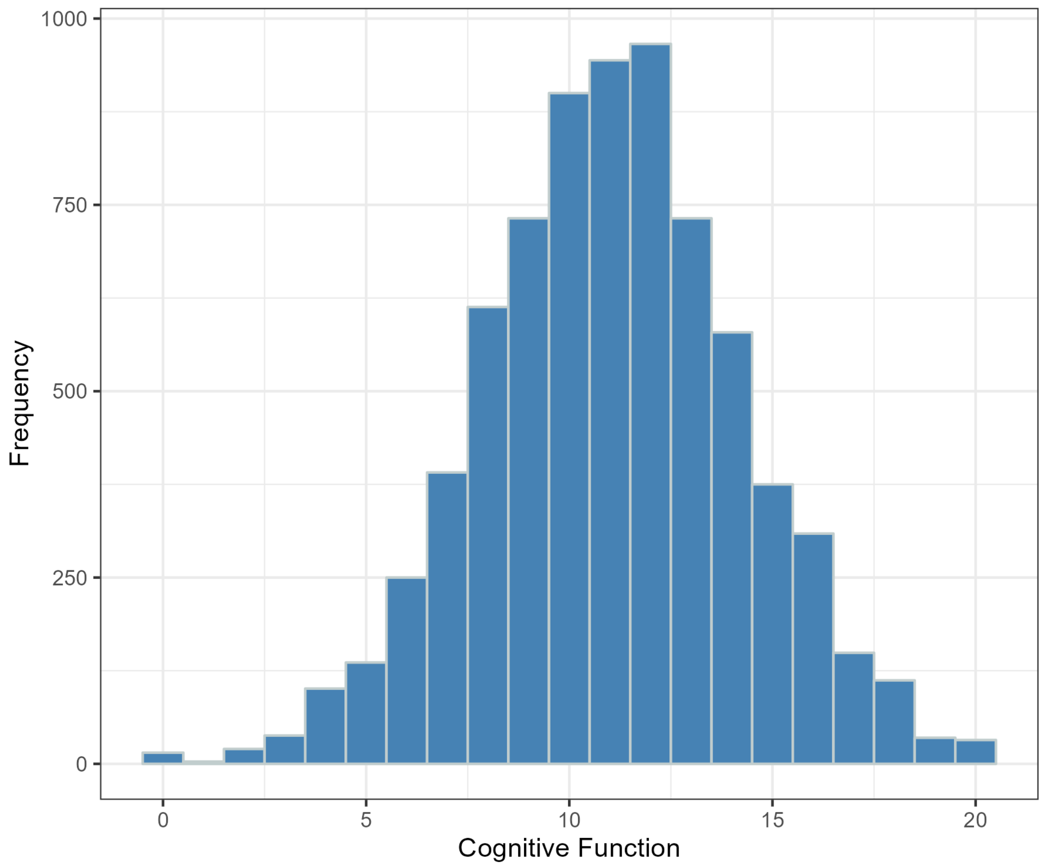

FIGURE B. DISTRIBUTION OF COGNITIVE FUNCTION

G. Partial Dependence Plot for Continuous Variables

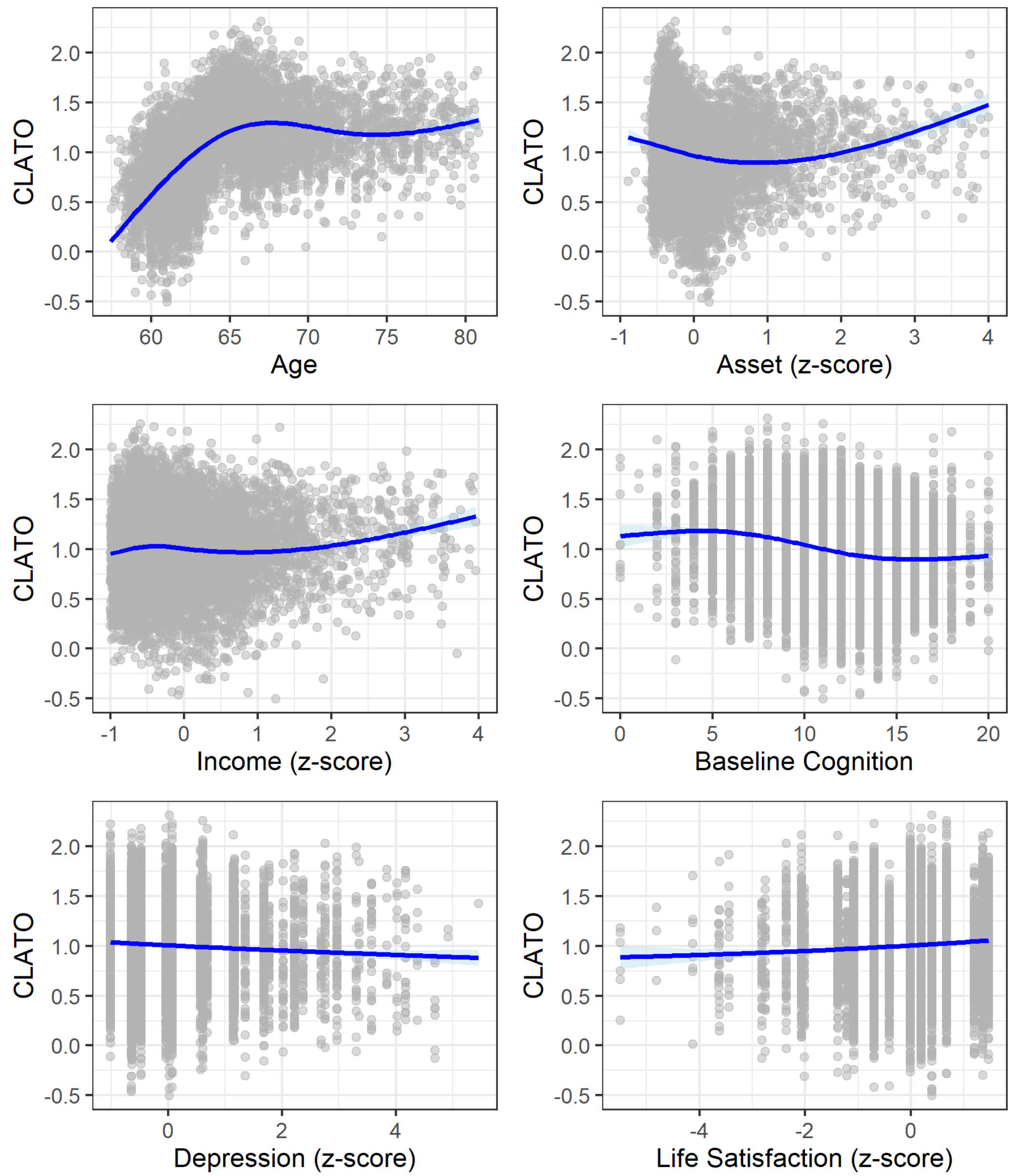

FIGURE C. PARTIAL DEPENDENCE PLOT FOR CONTINUOUS VARIABLES

Notes: CLATO stands for the conditional local average treatment effect on the overlap population.

H. Details in Monetary Cost Estimation

TABLE E—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.

Notes: RAVLT stands for Rey Auditory Verbal Learning Test, CI denotes confidence interval, HRS stands for the Health and Retirement Study, and ORA stands for official retirement age.

TABLE F—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.

Notes: RAVLT stands for Rey Auditory Verbal Learning Test, CI denotes confidence interval, ELSA stands for the English Longitudinal Study on Ageing, and ORA stands for official retirement age.