MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2020-013 I April 2020
https://doi.org/10.4054/MPIDR-WP-2020-013

Does postponing retirement affect cognitive function? A counterfactual experiment to disentangle life course risk factors

Jo Mhairi Hale I hale@demogr.mpg.de
Maarten J. Bijlsma I bijlsma@demogr.mpg.de
Angelo Lorenti I lorenti@demogr.mpg.de

This working paper has been approved for release by: Kieron Barclay (barclay@demogr.mpg.de), Deputy Head of the Laboratory of Population Health.

Does postponing retirement affect cognitive function?

A counterfactual experiment to disentangle life course risk factors

Jo Mhairi Hale[1,2]

Maarten J. Bijlsma[2]

Angelo Lorenti[2]

[1] University of St Andrews, Scotland

[2] Max Planck Institute for Demographic Research, Rostock, Germany

Abstract

Life-course sociodemographic and behavioral factors affect later-life cognitive function. Some evidence suggests that contemporaneous labor force participation also affects cognitive function; however, it is unclear whether it is employment itself or endogenous factors related to individuals' likelihood of employment that protects against cognitive decline. We exploit innovations in counterfactual causal inference to disentangle the effect of postponing retirement on later-life cognitive function from the effects of other life-course factors. With the U.S. Health and Retirement Study (1996-2014, n=20,469), we use the parametric g-formula to estimate the population-averaged effect (PAE) of postponing retirement to age 67, the average treatment on the treated (ATT), the moderating effect of gender, education, and occupation, and the mediating effect via depressive symptoms and comorbidities. We find that postponing retirement is protective against cognitive decline, accounting for other life-course factors (age 67 PAE: 0.34, 95% confidence interval (CI): 0.20,0.47; ATT: 0.43, 95% CI: 0.26,0.60). The extent of the protective effect depends on subgroup, with the highest educated experiencing the greatest reduction in cognitive decline (age 67 ATT: 50%, 95% CI: 32%,71%). By using innovative models that better reflect the empirical reality of interconnected life-course processes, this work makes progress in understanding how retirement affects cognitive function.

Keywords: cognitive impairment, dementia, labor force participation, retirement, life course

Does postponing retirement affect cognitive function?

A counterfactual experiment to disentangle life course risk factors

Introduction

Concerns about how lengthening life expectancies will affect health care and pension systems have led the U.S. government, like other high-income countries, to postpone the statutory retirement age for more recent cohorts. There may be a fortuitous unintended consequence of postponed retirement. Evidence is accumulating that cognitive engagement is associated with better cognitive function, which would imply that sustained participation in the labor force may be protective against cognitive decline (Adam et al. 2013; Bonsang et al. 2012; Kuiper et al. 2015; Meng et al. 2017; Mosca and Wright 2018; Reed et al. 2011; Roberts et al. 2011; Rohwedder and Willis 2010; Sharp 2017). However, because risk factors accumulate and interact over the life course to affect cognitive function, identifying and quantifying the effect of retirement on cognitive decline has proved elusive (Glymour and Manly 2008). In fact, research often focuses exclusively on single characteristics, such as work, education, or race/ethnicity, not accounting for how life-course factors are dynamically interconnected (Cho et al. 2013; Collins 2015; Diez Roux 2012).

In this paper we focus on retirement, but take into consideration race/ethnicity, gender, early-life socioeconomic status (SES), educational and occupational attainment, health behaviors, and mental and physical health. Understanding retirement's effect on cognitive function may help us anticipate trajectories of cognitive decline for more recent cohorts. These more recent cohorts will retire at older ages than earlier cohorts who had younger statutory retirement ages. For cohorts born in or after 1960, the U.S. federal government increased the full retirement age to 67. The statutory retirement age is important because most people retire at the

statutory early or normal retirement age (Behaghel and Blau 2012). We predict that an unintended consequence of this policy change may be, at the individual and population level, better cognitive function into older ages for these more recent cohorts. In other words, we hypothesize that postponed retirement will be protective against cognitive decline, accounting also for risk factors that vary across cohorts.

There are, however, some methodological challenges related to understanding life course risk factors for cognitive impairment that have prevented conclusive research on this topic. Biased estimates can be a result of: 1) dynamic two-way pathways among predictor and outcome, e.g., if retirement and cognitive function affect each other, 2) interactions between exposures and time-varying mediators, e.g., if retirement's effect on cognitive function differs by depressive symptoms, and 3) mediator-outcome confounding, e.g., if depressive symptoms affect cognitive function and retirement, but is itself affected by both (Kuiper et al. 2015; Sharp 2017).

The parametric g-formula offers a solution to these methodological issues. The g-formula is an innovative statistical approach which enables analysis of time-varying processes, while allowing for selection, reverse causality, and mediation (Bijlsma and Wilson 2020; Keil et al. 2014; Vanderweele and Tchetgen Tchetgen 2017; Wang and Arah 2015). Through a series of simulation steps, the parametric g-formula starts from the conditional distributions of the outcome (labor force participation) within levels of treatment (age at retirement), confounders (e.g., education), and mediators (e.g., depressive symptoms), thus approximating a standardized (with respect to the confounders and mediators) distribution of the outcome. As such, the g-formula is a statistically flexible approach that allows us to examine the interdependent influences of life-course processes – such as education, partnership, health, and labor force participation – on later-life cognitive function, irrespective of the functional form that the

relationships of mutual influence may take (De Stavola et al. 2015; Moffitt 2005; Schoeni et al. 2018; VanderWeele et al. 2014). An additional benefit of the g-formula is that while it models relationships at the individual level, it estimates population-averaged effects through hypothetical interventions, which is important from a demographic or public health policy perspective (Bijlsma, Tarkiainen, et al. 2017; Lin et al. 2017; VanderWeele and Tchetgen Tchetgen 2017; Wang and Arah 2015).

With data from the Health and Retirement Study (HRS) (University of Michigan 2017), we first use the parametric g-formula to estimate the causal effect on later-life cognitive function of postponing retirement to age 67, accounting for time-variant and invariant sociodemographic, behavioral, and health risk factors (Keil et al. 2014; Robins 1986). Second, we conduct moderation analyses to identify whether the effect of postponed retirement on cognitive function differs by gender, educational attainment, and/or occupational attainment. Third, we conduct mediation analyses to test whether depressive symptoms or health operate as mechanisms linking retirement and cognitive function.

Background

There is a growing body of research showing later-life cognitive function is affected by sociodemographic, behavioral, and health factors throughout the life course. A factor that is proximate to later-life cognitive function and yet for which evidence remains uncertain is labor force participation. Labor force participation is associated with demographic and life-course factors (e.g., educational and occupational attainment), as well as contemporaneous factors such as wealth, health behaviors, and health status. Traditional regression analysis has been unable to identify the salience of life-course risk factors because of a set of interrelated methodological barriers, variously expressed as longitudinal interdependence, intermediate and time-varying

confounding, as well as the challenges of analyzing intermediate variables and outcome variables with non-linear functional forms (Adler et al. 2012; Bijlsma, Tarkiainen, et al. 2017; Jones et al. 2011; Mehta and Preston 2016).

Researchers have employed a variety of solutions to try and address these methodological challenges. From studies using cross-national variation in the statutory retirement age and instrumental variable techniques, there is mixed evidence that retirement is negatively associated with health, broadly, or cognitive function, specifically (Adam et al. 2013; Bonsang et al. 2012; Celidoni et al. 2017; Clouston and Denier 2017; Coe et al. 2012; Coe and Zamarro 2011; Denier et al. 2017; Mazzonna and Peracchi 2012; Rohwedder and Willis 2010). For example, using the HRS and ordinary least squares regression, Coe and colleagues (2012) find retirement is associated with lower cognitive function. However, when they use time between interview and offer of early retirement as an instrument, they find no negative causal effect of retirement on cognitive function for white-collar workers and a slight positive effect for blue-collar workers. In contrast, Clouston and Denier (2017) and Bonsang and colleagues (2012), respectively, use longitudinal regression discontinuity models and eligibility for social security as an instrument. Both studies do find a benefit of continued labor force participation on cognitive function. Moreover, this research focuses on the short-term effects of retirement on cognitive functioning; whereas, we evaluate how retirement affects cognitive functioning in the longer term, as well as pathways of influence on cognitive functioning (e.g., depressive symptoms). In short, evidence on the association between retirement and cognitive function remains inconclusive, and methodological problems are the primary barrier.

We use a different modeling strategy to overcome these methodological challenges. The g-formula significantly decreases the bias likely when using standard regression analysis in the

presence of longitudinal interdependence. It allows us to disentangle mechanisms (time-variant or invariant, of any parametric functional form or distribution) across the life course (Bijlsma, Daniel, et al. 2017; Bijlsma, Tarkiainen, et al. 2017; Daniel et al. 2013; VanderWeele and Tchetgen Tchetgen 2017). The g-formula approach estimates effects of the determinant on the outcome through simulating hypothetical interventions, such as the effect of postponed retirement on cognitive function (Bijlsma, Tarkiainen, et al. 2017). This can answer the question, for example, to what extent the new statutory full retirement age of 67 might delay cognitive decline compared with retiring at younger ages. It is worth mentioning that with our approach we account also for individuals reentering the labor market after retirement, while other studies cannot. Evidence suggests that only 37% of U.S. workers actually retire completely from a full-time job (Hudomiet et al. 2018). In addition, the g-formula approach enables the estimation of both population-averaged effects and the effect of the treatment on the treated (Lin et al. 2017; Wang and Arah 2015). In this case, the latter would be the effect of postponed retirement only on those who actually retired prior to age 67, versus on the population more widely.

Furthermore, because postponed retirement's effect on cognitive function may differ among subpopulations, we examine whether there are differences in the effect of postponing retirement on cognitive function by gender, educational attainment, and occupational attainment. Gendered differences in how identity is tied to paid employment and in propensity to engage socially may contribute to gender moderating retirement's effect on cognitive decline (Barnes and Parry 2004; Kuiper et al. 2015). We, therefore, hypothesize that postponing retirement may benefit men more than women. The effect of postponing retirement on cognitive function may also depend on level of educational attainment. Those with higher education are likely to be in more cognitively stimulating, complex occupations, and those with lower education may be in

more physically demanding occupations. As individuals employed in intellectually-stimulating occupations are over-represented among the highly educated, we expect this combination could mean postponing retirement is more beneficial to the higher educated (Potter et al. 2008). Relatedly, evidence suggests that level of job complexity or intellectual demands are associated with later-life cognitive function (Andel et al. 2005; Fisher et al. 2014; Potter et al. 2008; Staudinger et al. 2016). If intellectually-demanding work is associated with better cognitive function, those who work in more monotonous jobs may not experience as much of a benefit from postponing retirement. Health insults and socioeconomic stressors related to "bad jobs" (Kalleberg et al. 2000) may also hinder benefits of ongoing labor force participation. At the same time, it is possible that those in non-professional occupations may benefit from postponed retirement because the working environment, per se, is an important stimulus, especially if retirees have few cognitively-engaging non-work activities (Gow et al. 2016; Reed et al. 2011). These countervailing forces lead us to cautiously hypothesize that postponing retirement could be more beneficial for those with white-collar jobs.

We also investigate two mechanisms through which retirement may affect cognitive decline. Those who retire might subsequently experience an increase in depressive symptoms and/or comorbidities that partially explain the effect of retirement on cognitive decline. Thus, we examine whether depressive symptoms or comorbidities, which are associated with both labor force participation and cognitive function (Calvo et al. 2013; Fonseca et al. 2014; González et al. 2008; Virtanen et al. 2015), operate as mediators.

Hypotheses

Based on this growing body of literature and fully exploiting the flexibility of the g-formula, we derive the following hypotheses:

Hypothesis 1: Compared with retiring between age 50 and 66, postponing retirement until age 67 or older will be protective against cognitive decline, accounting for gender, race/ethnicity, birth cohort, early-life SES, educational and occupational attainment, partnership status, exercise, alcohol consumption, depressive symptoms, and comorbidities.

Moderation analysis

Hypothesis 2: Accounting for all above covariates, postponing retirement to age 67 will be differentially protective against cognitive decline for:

a) men more than women,
b) higher educated more than those with less than a high school education,
c) professional occupations more than those in non-professional occupations,

Mediation analysis

Hypothesis 3: Depressive symptoms and comorbidities act as mechanisms, mediating the effect of postponing retirement on later-life cognitive function.

Methods

Dataset

The Health and Retirement Study (HRS) (1992-ongoing) is a longitudinal, nationally-representative, biennial survey of U.S. residents age 55 and over and their spouses (regardless of age) that includes many sociodemographic, wealth, and health measures. The University of Michigan conducts the HRS, which is sponsored by the National Institute on Aging (grant number NIA U01AG009740). We use RAND Version P of the HRS (RAND Center for the Study of Aging 2017). The HRS uses a modified version of the Telephone Interview for Cognitive Status (TICS-M) that is sensitive to pathological cognitive decline (Fong et al. 2009; Karlamangla et al. 2009). We use the University of Michigan Survey Research Center's imputed

values for the TICS-M measures (Fisher et al. 2017).

The HRS includes retrospective data on early-life environment and educational attainment, as well as biennial data on health behaviors, health, and labor force participation. We use data from all waves in terms of collecting retrospective data, but focus on the period 1996 to 2014—the years for which consistent data is available for the cognitive function measures we include. We extracted individuals who were age 55 to 75 years old, who: had self-responses for TICS-M (we analyze cognitive function as a continuous outcome, for which proxy responses are not adequate), have participated in the labor market at some point in the 1996-2014 period, were not retired prior to 1996, and have non-missingness on other covariates (less than 1% missing on any covariate) (Bonsang et al. 2012). The final analytical sample is 96,918 observations from 20,469 individuals.

Key Study Measures

Cognitive Function. We extract a subset of questions from the HRS TICS-M that represent fluid intelligence—a composite of cognitive domains reflective of neurophysiological health (Akshoomoff et al. 2013; Ghisletta et al. 2012; Horn 1982). We select immediate (0-10 points) and delayed word recall (0-10 points), serial 7s (counting backward from 100 by sevens) (0-5 points), and counting backward from twenty (0-2 points). The range is 0-27, where higher values represent better cognitive function.

Primary exposure

The primary (time-varying) exposure is employment status. Employment status is a categorical (multinomial) variable that indicates if an individual is full-time employed, part-time employed/part-time retired, unemployed, disabled, or retired. This measure is not based on hours worked, but the self-reported category. On average, those full-time working report at least 40

hours per week, and those who report being part-time employed or part-time retired work about 25 hours per week.

Time-varying covariates

Covariates include: *partnership status* (partnered, separated/divorced/spouse absent, never married, widowed), *weekly exercise* (1=exercise once or more per week), *alcohol consumption* (abstinent/rare, light, moderate, heavier), *depressive symptoms* (Center for Epidemiological Studies-Depression, CES-D, 0, 1, 2-4, or 5-8 symptoms), and *comorbidities* (0-4, diagnoses of stroke, diabetes, heart condition, and/or high blood pressure/hypertension).

Time-invariant covariates

Birth cohort follows the HRS cohort structure (AHEAD 1919-1923, Children of the Depression Era 1924-1930, HRS 1931-1941, Warbabies 1942-1947, Early Babyboomers 1948-1953, Mid Babyboomers 1954-1959). HRS reports a binary *Gender* variable (1=Women). *Race/Ethnicity* is Non-Hispanic White, African American/Black Hispanic, Non-Black Hispanic, and Non-Hispanic Other (henceforth White, Black, Latinx, Other). *Age* is age in years. To provide a more comprehensive picture of the early-life environment, *Early-SES* includes self-reported childhood SES, childhood health, parents' education, father did not contribute economically (unemployed, absent, dead), father's lower-status occupation, childhood family moved due to financial hardship, and/or childhood family borrowed money due to financial hardship (0, 1, 2-5, or 6-7 adversities). *Educational attainment* is defined as less than high school/general equivalency diploma (GED), high school diploma, and some college or higher. *Longest job* ever held is categorized as professional or non-professional (the latter includes sales, administrative, service, manual, farms, forestry, and fishing). *Wealth* is a RAND-generated measure that includes household income, assets, and debts, which we average over the study

period to get a time-invariant measure of household wealth (in debt; 0-$49,999; $50,000-$199,999; $200,000-$499,999; $500,000-$999,999; $1 million or more).

Analytical strategy

In short, the g-formula approach is implemented following four steps. First, we construct a causal directed acyclic graph (DAG, Figure 1), a diagram that portrays the interrelationships among the variables we will model. Second, using the DAG as guidance, we estimate a series of multivariable models for the intermediate and outcome variables. Time-varying variables at age *a* are allowed to be affected by all time-invariant variables and, to limit assumptions on causality within a calendar year, by all time-varying variables in the previous year. We model categorical variables using multinomial logistic regression models and continuous variables with linear regression models. Third, we define intervention scenarios (e.g., what happens if people retire at older ages?). Fourth, using empirical observations at age 55, the estimated multivariable models, and following the DAG, we simulate an approximation of the empirical data (the natural course scenario) as well as an approximation of the sample under a hypothetical intervention (the intervention scenario). More detailed information on the third and fourth steps is provided below.

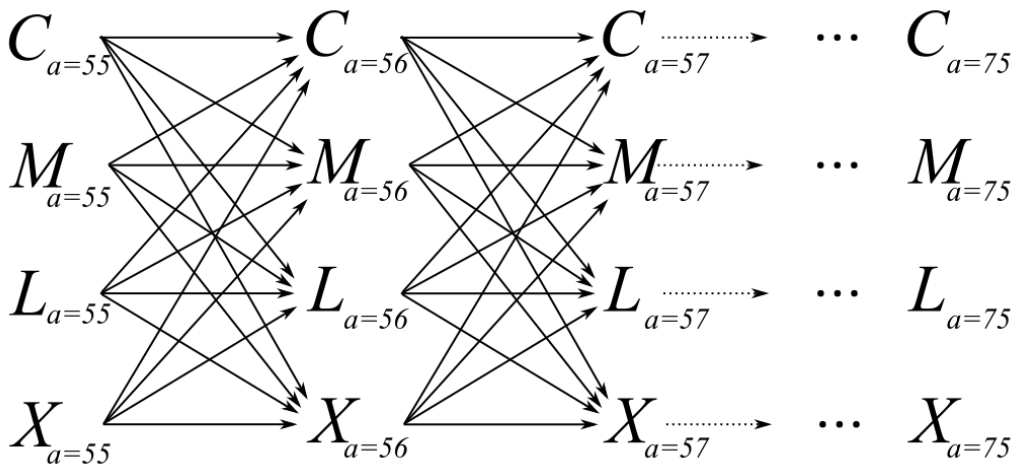

Figure 1 Simplified directed acyclic graph (DAG) showing the single-year cross-lagged structure whereby cognitive function (C), mediating factors (M), labor force participation (L) and time-varying confounders (X) are associated across age (a) 55 to 75. For simplicity, the DAG does not show time-invariant control variables, but these are included in all models.

Natural course vs. intervention

To test Hypothesis 1, we contrast a scenario in which all observations were as empirically observed (the "natural course scenario") with a scenario in which all individuals who retired before age 67 instead maintained employment until at least age 67 (the "intervention scenario"), after which they can retire following observed retirement risks for individuals with their covariate distribution. This is accomplished in the simulation step of the g-formula, including a 500-iteration bootstrap to produce standard errors and confidence intervals (Efron and Tibshirani 1994; Jain et al. 2016; Keil et al. 2014). In each iteration of the bootstrap, we randomly drew individuals with replacement from the data and re-estimated the earlier specified multivariable models on these data. Then, using the first observations of these individuals and the estimates from the multivariable models, we simulated their observations in the second year. For binomial variables, simulated values were drawn stochastically from a binomial distribution with the predicted mean based on the corresponding multivariable model; for multivariate variables this was done in the same manner, but using a multinomial distribution; and for the continuous

variable (cognitive function), this was done using a Gaussian distribution (with variance based on the prediction model's residual variance) (Robins 1986). From these simulated values, we simulate values in the third year, and so on, until the end of follow-up (Keil et al. 2014). We produce the estimates in the intervention scenario similarly, but whenever someone retired before or at age 67, we instead keep them employed. By taking the differences between the intervention scenario and the natural course scenario, we calculate the total effect of our intervention of postponing retirement to age 67 (Wang and Arah 2015).

Subgroup analysis: Gender, educational attainment, and occupational attainment

Hypothesis 2 is that postponing retirement to age 67 will be differentially protective against cognitive decline for men more than women, the higher educated more than the lower educated, and those in professional more than non-professional occupations. To test this, we allowed for interaction terms between labor force participation and gender, education, and occupation within the multivariable models. This allows for separate effects of employment and retirement on cognitive function by gender, education, and occupational group. Following the g-formula procedure explained above, we then save and compare results separately by subgroup.

Mediation analysis: depressive symptoms and comorbidities

To test Hypothesis 3, we perform mediation analyses in which we determine to what extent the effect of labor force participation on cognitive function is mediated through depressive symptoms and comorbidities. The direct effect of postponing retirement is determined by performing simulations that are identical to the intervention scenario (as in Hypothesis 1) for the total effect, with the exception that mediators of interest are kept at their natural course levels. Keeping the mediators at their natural course levels prevents the intervention from affecting the mediators, thus eliminating the part of the intervention effect that operates 'via' these mediators.

The indirect effect, the portion that does operate via the mediators of interest, is determined by subtracting this direct effect from the total effect. This allows us to examine to what extent the effect of postponing retirement operates through retirement's effect on depressive symptoms or comorbidities. See also Appendix III for more information on the mediation analysis.

Sensitivity Analyses

In the intervention scenarios presented above, after postponing retirement until age 67, individuals were "allowed" to retire following empirical expectations conditional on their covariate values. Using this scenario, any difference in cognitive function between the intervention and natural course scenarios that is evident after age 67 could be caused both by the enduring effect of postponed retirement at the individual level and by having a larger number of not-yet retired individuals after age 67 in the intervention scenario relative to the natural course.

Therefore, we produced an additional comparison where, in both the intervention and the natural course scenario, all individuals were hypothetically "forced" to retire at age 67, This comparison eliminates compositional differences in the number of retired individuals at age 67 and older; hence, only the enduring protective effect of postponing retirement at the individual level leads to the population-level differences in cognition after age 67. These results are in Appendix II.

Results

Table 1 shows descriptive statistics for the analytic sample. About 46% of the observations are in full- or part-time work, while almost 45% are in the retirement state. Retired and disabled individuals have the lowest cognitive function scores. Those who are retired are older and more likely to be women than those who are working. The persistence and interconnectedness of disadvantage over the life course is clear in that Blacks and Latinx are more likely to be

unemployed or disabled and less likely to be retired, and those with more early-life disadvantages and lower educational attainment are more likely to be disabled. Those working are more likely to exercise and have fewer comorbidities or depressive symptoms.

Table 1 Descriptive statistics for the analytic sample by labor force participation

	FT	PT	Retired	Unem-ployed	Disabled	NILF	Total
Variable	*Mean/%*	*Mean/%*	*Mean/%*	*Mean/%*	*Mean/%*	*Mean/%*	*Mean/%*
LFP distribution	29.3	16.9	44.6	2.1	2.3	4.8	100
Cognitive Function	16.9	16.6	15.3	15.8	13.5	15.6	16.0
#Cognitive Tests	5.9	6.3	6.2	5.2	5.5	6.3	6.1
Age	60.3	64.4	67.3	61.1	60.7	63.4	64.3
Female	48.1	60.1	56.0	49.9	64.2	93.7	56.0
			Rows total to 100%				
Race/ethnicity							n person-waves
White	28.8	17.7	45.9	1.7	1.6	4.4	74,791
Black	29.3	15.3	44.9	2.8	4.6	3.1	17,122
Latinx	31.6	14.2	35.9	3.6	4.0	10.6	9,820
Other	35.4	15.1	38.7	3.6	3.1	4.1	2,517
Early-Life Socioeconomic Status							
6-7	23.3	14.2	49.9	2.2	5.9	4.5	1,869
2-5	26.4	16.0	47.4	2.1	2.9	5.1	54,430
1	29.6	17.5	44.4	2.0	1.7	4.8	29,377
0	38.0	18.8	36.0	2.4	1.1	3.7	18,574
Educational Attainment							
Less than HS/GED	20.8	14.1	50.5	2.1	5.0	7.6	24,951
HS Diploma	28.6	16.9	45.8	2.0	1.8	4.8	50,828
Some College+	38.0	19.5	37.2	2.2	0.9	2.2	28,471
Longest Occupation							
Non-professional	27.1	16.2	45.7	2.2	3.0	5.9	71,320
Professional	34.1	18.5	42.3	1.9	0.9	2.3	32,930
Wealth Average							
In debt	26.8	13.4	38.9	5.9	9.3	5.7	4,138
$0-49K	27.0	14.3	44.3	3.0	5.6	5.7	20,688
$50-199K	30.8	16.0	44.6	2.0	1.7	4.9	31,860
$200-499K	29.5	17.9	45.8	1.7	0.9	4.2	25,307
$500-999K	28.6	18.7	47.0	1.3	0.5	4.0	13,846
$1mil+	31.0	22.7	40.5	0.9	0.4	4.6	8,411
Partnership Status							
Never Married	33.9	14.8	41.2	3.0	4.9	2.0	3,906
Married/Partnered	30.3	17.4	43.3	1.9	1.7	5.3	72,273
Sep/Divorced/Absent	33.7	15.9	40.1	3.3	4.2	3.0	15,092
Widowed	17.1	16.0	57.9	1.4	2.7	4.9	12,979
Exercise							
Weekly or less	27.0	14.6	47.8	2.0	3.5	5.2	50,726

More than weekly	31.5	19.2	41.6	2.2	1.2	4.4	53,524
Alcohol consumption							
Abstinent/rare	26.7	15.9	46.6	2.0	2.9	5.8	66,957
Light	31.1	19.6	43.4	1.7	1.0	3.2	15,673
Moderate	36.1	19.1	38.6	2.2	1.0	3.0	12,611
Heavier	36.0	16.7	39.8	3.4	1.8	2.3	9,009
CESD							
0	33.1	19.2	41.4	1.7	0.6	3.9	49,250
1	30.1	17.1	44.2	2.1	1.7	4.8	22,858
2-4	24.8	14.1	49.3	2.4	3.9	5.5	22,241
5-8	18.4	11.4	50.7	3.7	8.5	7.3	9,901
Comorbidity Index							
None	37.3	19.1	34.5	2.3	1.4	5.4	36,907
One	29.4	17.4	44.7	2.1	1.9	4.5	38,696
Two	21.6	14.6	53.9	2.0	3.3	4.5	20,938
Three	12.4	11.1	65.5	1.5	5.9	3.7	6,645
Four	5.8	6.2	74.6	1.1	8.6	3.7	1,064

Figure displays the population-averaged effect (PAE) of postponing retirement to at least age 67. Even at the population level (i.e., where even people who did not retire prior to age 67 are in the denominator), there is a positive effect for both women and men of postponing retirement until age 67 or older, accounting for all time-invariant characteristics (cohort, race/ethnicity, early-life SES, and educational and occupational attainment) and accounting for the bidirectional association between labor force participation and time-varying factors (partnership status, exercise, alcohol consumption, depressive symptoms, and comorbidities) (Figure). This intervention scenario compared to the natural course consistently shows a positive effect throughout the age range, including a positive effect after age 67 up until at least age 74 for both men and women (Figure 2 and Appendix I). However, as explained above, we cannot attribute the extent of this positive effect only to an enduring protective effect of postponed retirement. The effect could be driven also by a larger number of individuals who continue to work at older ages, as this intervention scenario "allows" individuals to retire older than 67.

Therefore, in sensitivity analysis, we also examine scenarios whereby instead of "allowing" respondents in the intervention scenario to exit the labor force in phases, we "force"

everyone to retire at age 67. In this scenario comparison, any effect at age 67 or older is evidence of an enduring protective effect of postponing retirement. Even in this forced-retirement scenario, the protective effect lasts at least five years post retirement, up to at least age 72 for both men and women (Appendix II). This protective effect operates not because labor force participation improves cognitive function, but because in the natural course scenario, those who retire younger than 67 experience faster cognitive decline. This is strong evidence for Hypothesis 1.

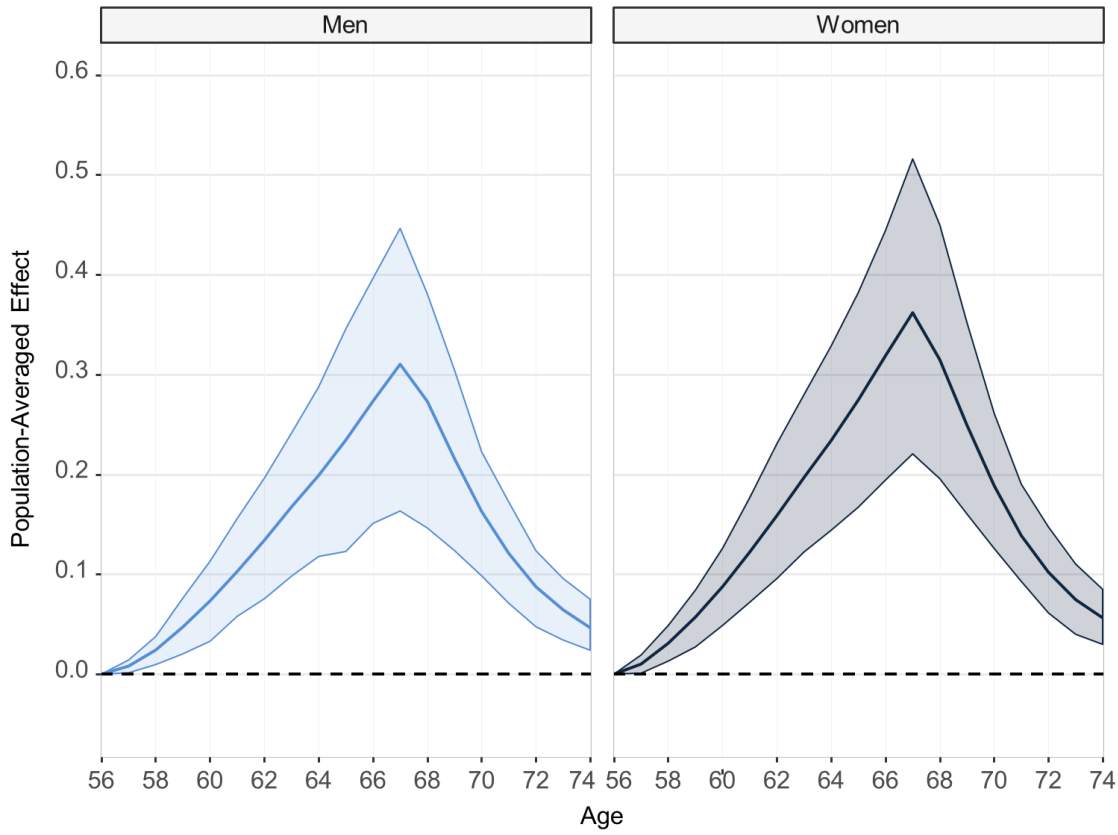

Figure 2 Population-averaged effect (PAE)—the difference in cognitive functioning between the natural course scenario and the intervention where retirement is postponed until at least age 67.

In interpreting the meaning of the population-averaged effect, it is informative to compare this protective effect with the number of points lost in the natural course scenario. The average age at retirement for those who retire before age 67 is age 61, which does not differ by subgroup

(by gender, race/ethnicity, education, or occupational attainment). From age 61 to age 67, the average change in cognition in all subgroups is approximately 1 point on the 0-27 scale, ranging from least lost for those with less than high school (0.92) to most lost for those in professional occupations (1.06). Delaying retirement to age 67 allows men to retain a cognition score that is 0.31 (95% CI: 0.16, 0.45) points higher than if their retirement had not been delayed, and for women this is 0.36 points higher (95% CI: 0.22, 0.52). Therefore, relative to the decline over the age 61 to 67 period, this represents approximately a one-third reduction in cognitive decline over the relevant time period. Note that only those who retire prior to age 67, approximately 63% of the sample, are affected by the intervention. Individuals who work up to age 67 are not affected by the intervention, but do contribute to the denominator of the population-averaged effect. In other words, they dilute the population-averaged effect estimate.

Whereas above we present the population-averaged effect, for Hypothesis 2, we present the average effect of the treatment on the treated (ATT); that is, the effect of postponed retirement on only those who actually do retire prior to age 67. Figure shows that, indeed, the effect on the individual of postponed retirement is larger than the population-averaged effect (men: 0.42, 95% CI: 0.22, 0.59; women: 0.44, 95% CI: 0.27, 0.63). Again, the protective effect lasts well beyond age 67 (Appendix II). However, we find no support for Hypothesis 2a that men differentially benefit from postponed retirement compared with women.

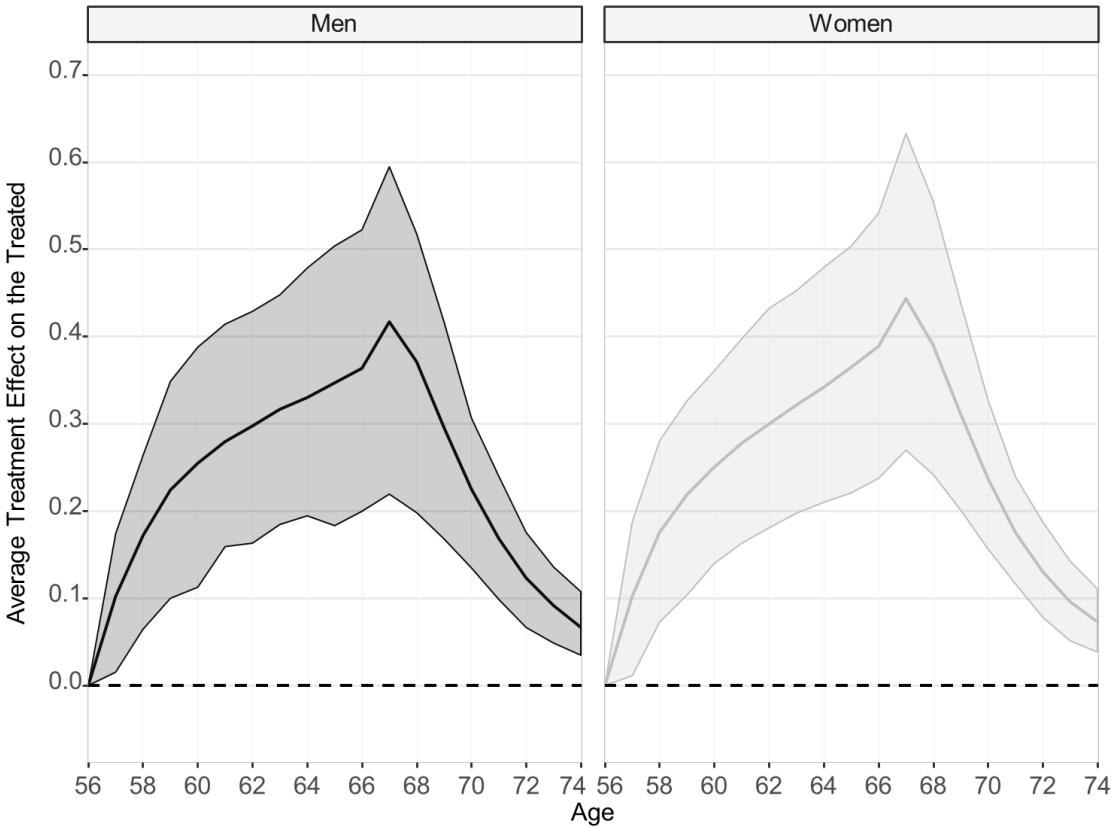

Figure 3 Average treatment effect on the treated by gender— the difference in cognitive function score between the natural course and intervention scenarios only for those who retired prior to age 67.

Similarly, Figure presents the average treatment effect on the treated for those with less than high school/GED, a high school diploma, or some college or more. Those with higher educational attainment have slightly higher point estimates across all ages than those with lower educational attainment (<HS: 0.31, 95% CI: 0.09, 0.48; HS: 0.46, 95% CI: 0.26, 0.64; Some college or higher: 0.50, 95% CI: 0.32, 0.71); however, we find no clear evidence for our Hypothesis 2b that the higher educated would differentially benefit. The effect size for those with at least some college is particularly substantial, suggesting postponed retirement is associated with an approximately 50% reduction in cognitive decline.

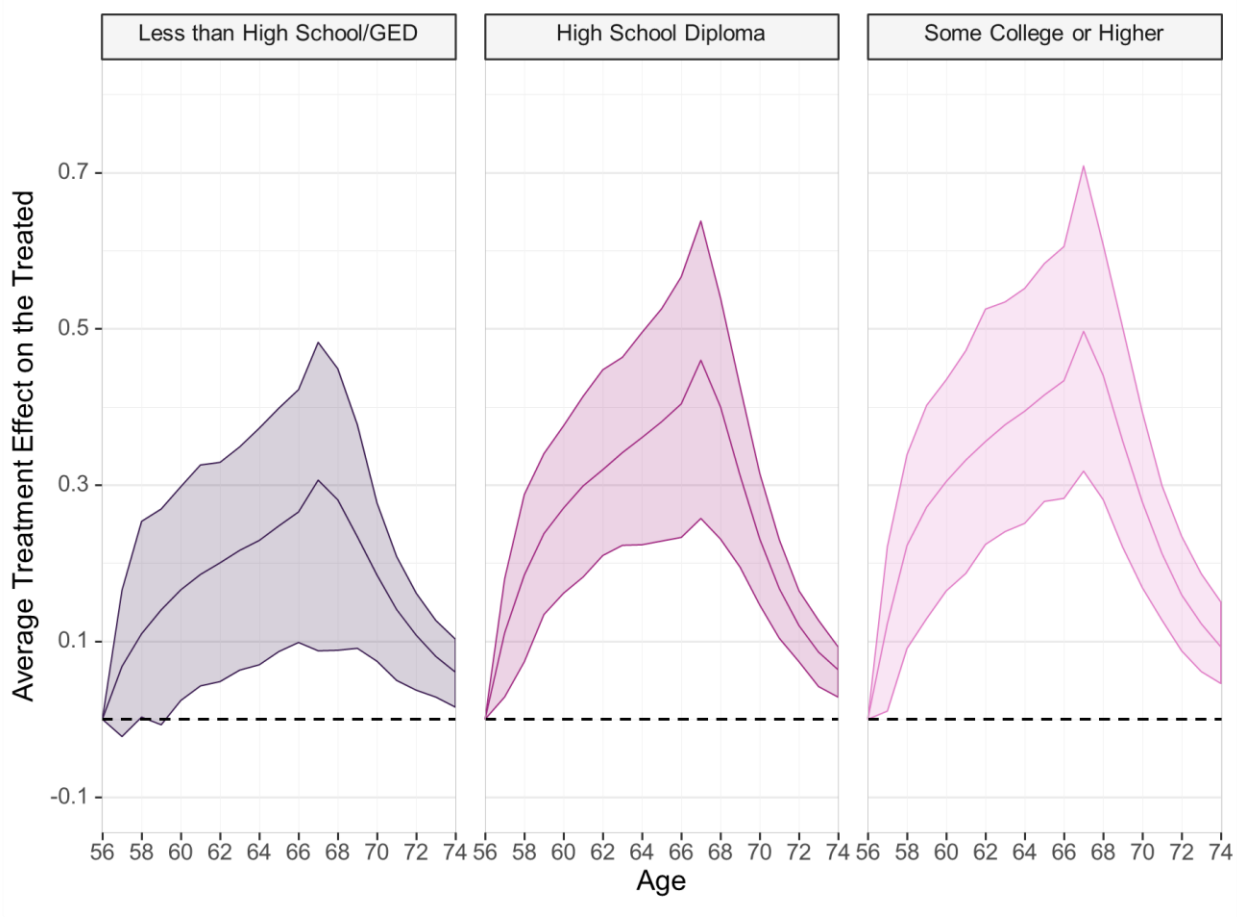

Figure 4 Average treatment effect on the treated by educational attainment—the difference in cognitive function score between the natural course and intervention scenarios only for those who retired prior to age 67.

Both those in non-professional (0.45, 95% CI: 0.27, 0.64) and in professional occupations (0.37, 95% CI: 0.16, 0.57) appear to benefit from postponed retirement, with a slightly greater, but non-significant benefit to the former (Figure 5). Thus, our findings do not support Hypothesis 2c.

Figure 5 Average treatment effect on the treated (ATT) by occupational attainment—the difference in cognitive function score between the natural course and intervention scenarios only for those who retired prior to age 67.

We next test Hypothesis 3, where we conduct mediation analyses to identify to what extent depressive symptoms or comorbidities operate as mechanisms through which retirement affects cognitive function. We find no evidence for Hypothesis 3 that depressive symptoms or comorbidities significantly mediate the effect of retirement on cognitive function (Appendix III).

Discussion

Interdependent life-course processes influence cognitive function in terms of "achieved" cognition, as well as rate of decline. However, there are critical gaps in our knowledge about the modifiable factors that may be protective against later-life cognitive decline, a shortcoming

related both to a lack of research that takes into consideration that life-course factors are dynamically interconnected and the related methodological barriers. The g-formula is a more flexible modeling strategy that better approaches the empirical reality of the life course's influence on later-life cognitive health.

Accounting for demographic and early-life factors, as well as the longitudinal interdependence between educational and occupational attainment, labor force participation, and health, we find evidence for Hypothesis 1 that postponing retirement to age 67 provides an insulative effect against cognitive decline. Indeed, even the population-averaged effect of the intervention shows a 30-34% reduction, for men and women, respectively, in cognitive decline associated with remaining employed compared with retiring younger than age 67. The effect is related to a slowed rate of cognitive decline versus a "boost" in cognitive function. The protective effect appears to hold regardless of gender, educational attainment, or occupational attainment, thus we find no clear evidence for Hypothesis 2 that certain subgroups would differentially benefit from postponed retirement. We hypothesized that a mechanism through which retirement may affect cognitive function may be related to experiencing some level of depression and/or health insults as a result of retirement. We therefore examined whether depressive symptoms or comorbidities explained any of the effect of retirement on cognitive function. We did not find evidence that either of these explained much of the association between retirement and cognitive function.

Limitations

The causal claims of this study rely on three fundamental assumptions: positivity, consistency and exchangeability (Greenland and Robins 2009; Petersen et al. 2012; Rehkopf et al. 2016). The positivity assumption requires that individuals who were hypothetically intervened

on had in fact a non-zero chance of receiving such an intervention in the real world. In our study, this requires that individuals who hypothetically had their retirement postponed could in fact be employed. Our intervention was not performed on those individuals who were disabled or out of the labor force. However, there may have been other reasons for individuals to retire for which we do not have information. Since the intervention would not have been possible for these individuals, the true population-averaged effect would be smaller. Nevertheless, this issue will not have a substantial effect on the estimated treatment effect for the treated.

The consistency assumption requires that the hypothetical intervention of interest is well-defined and that the variable representing it in the dataset corresponds to this definition. In our study, our hypothetical intervention was a postponement of retirement to age 67 for those individuals who empirically retired before age 67. The variable used to hypothetically implement this intervention was a variable indicating employment status, and its effect is, by definition, drawn from those individuals who are employed at each age (relative to those who are not). Individuals who continue working at certain ages may do so for a variety of reasons, most prominently for health and financial reasons. Since we use information from these individuals, our hypothetical intervention does not represent a forced intervention, but a scenario that represents individuals choosing to work longer of their own accord (for any number of reasons). An important limitation is that some individuals have chosen to retire because of causes that are not represented (to the same degree) in the population of individuals who continue to work. This brings us to the exchangeability assumption.

Applied to our study, the exchangeability assumption requires that individuals who are employed are comparable, in terms of factors that affect cognitive function, to individuals who are retired, conditional on the measured covariates. For example, individuals may choose to

retire because their health has deteriorated, and they are no longer able to perform full or part-time work. If this deteriorated health has also affected their cognitive function, then this effect of health on cognitive function – if it has not been adjusted for in the study – will bias the association between employment and cognitive function away from the causal effect of employment *on* cognitive function. In our study, we adjust for a large number of potential confounding variables; these include gender, race/ethnicity, early-life disadvantage, education, occupation, weekly exercise, alcohol consumption, depression score, and comorbidities. However, it is likely that the models have not adjusted for some important confounding variables or that those that are measured do not perfectly capture salient matters (an issue known as residual confounding). Hence, although the aim of this study is to approximate a causal effect, we acknowledge that various biases likely still remain. A main limitation is that our ability to test for alternative mechanisms is limited by the data. Work conditions, tasks, and the interplay of market and non-market activities are factors that can contribute to the rate of cognitive decline. Indeed, for analytical reasons (power) and data constraints (no information, e.g., on tasks), our measure for occupation is particularly limited. Having the possibility to check specific work-related factors would help to understand why we find that working longer is beneficial. This is an important line for future research.

Relatedly, other research suggests that socially- or cognitively-engaging activities that are not recorded as labor market participation, e.g., grandparenting or volunteering, may also be protective (Gow et al. 2016; Reed et al. 2011). It is plausible that the relationship between retirement and cognitive function is driven by changes in engagement that we were unable to identify with the data available across the 1996-2014 timespan of the HRS. In order to test this

hypothesis, future research should exploit other datasets that include more detailed measures of social activities, such as the National Social life, Health, and Aging Project (NSHAP).

Conclusion

This work has significant implications. Longer life expectancies and population aging have motivated many high-income countries to postpone the statutory retirement age for more recent cohorts. Evidence suggests that labor market participation may be protective against cognitive decline, but because labor market participation is dynamically interconnected with other life-course factors that influence cognitive function, results are inconclusive. Using advanced counterfactual modeling significantly decreases the bias from issues involving longitudinal interdependence, compared with standard modeling approaches. This more flexible modeling strategy allows us to test the effect sizes of hypothetical intervention scenarios. In this case, the U.S. government already has raised the full retirement age for successive cohorts, and for those born after 1960, that age is 67. Our findings suggest that these more recent cohorts may, indeed, enjoy better cognitive function at older ages than their counterparts from earlier cohorts who were more likely to retire at younger ages.

Works cited

Adam, S., Bonsang, E., Grotz, C., & Perelman, S. (2013). Occupational activity and cognitive reserve: implications in terms of prevention of cognitive aging and Alzheimer's disease. *Clin Interv Aging, 8*, 377–390. doi:10.2147/cia.s39921

Adler, N. E., Bush, N. R., & Pantell, M. S. (2012). Rigor, vigor, and the study of health disparities. *Proc Natl Acad Sci USA, 109 Suppl*(Supplement_2), 17154–17159. doi:10.1073/pnas.1121399109

Akshoomoff, N., Beaumont, J. L., Bauer, P. J., Dikmen, S. S., Gershon, R. C., Mungas, D., et al. (2013). NIH toolbox cognition battery (CB): Composite scores of crystallized, fluid, and overall cognition. *Monographs of the Society for Research in Child Development, 78*(4), 119–132. doi:10.1111/mono.12038

Andel, R., Crowe, M., Pedersen, N. L., Mortimer, J., Crimmins, E., Johansson, B., & Gatz, M. (2005). Complexity of work and risk of Alzheimer's disease: A population-based study of Swedish twins. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 60*(5), 251–258. doi:10.1093/geronb/60.5.P251

Barnes, H., & Parry, J. (2004). Renegotiating identity and relationships: men and women's adjustments to retirement. *Ageing and Society, 24*(2), 213–233. doi:DOI: 10.1017/S0144686X0300148X

Behaghel, L., & Blau, D. M. (2012). Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age. *American Economic Journal: Economic Policy, 4*(4), 41–67. doi:10.1257/pol.4.4.41

Bijlsma, M. J., Daniel, R. M., Janssen, F., & De Stavola, B. L. (2017). An Assessment and Extension of the Mechanism-Based Approach to the Identification of Age-Period-Cohort

Models. *Demography*, *54*(2), 721–743. doi:10.1007/s13524-017-0562-6

Bijlsma, M. J., Tarkiainen, L., Myrskylä, M., & Martikainen, P. (2017). Unemployment and subsequent depression: A mediation analysis using the parametric G-formula. *Social Science and Medicine*, *194*(May), 142–150. doi:10.1016/j.socscimed.2017.10.011

Bijlsma, M. J., & Wilson, B. (2020). Modelling the socio-economic determinants of fertility: a mediation analysis using the parametric g-formula. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *183*(2), 493–513.

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

Calvo, E., Sarkisian, N., & Tamborini, C. R. (2013). Causal effects of retirement timing on subjective physical and emotional health. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *68*(1), 73–84. doi:10.1093/geronb/gbs097

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. doi:10.6103/SHARE.w3.100

Cho, S., Crenshaw, K. W., & Mccall, L. (2013). Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis. *Signs: Journal of Women in Culture and Society*, *38*(4), 785–810.

Clouston, S. A. P., & Denier, N. (2017). Mental retirement and health selection: Analyses from the U.S. Health and Retirement Study. *Social Science and Medicine*, *178*, 78–86. doi:10.1016/j.socscimed.2017.01.019

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.

doi:10.1002/hec.1771

Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of health economics*, *30*(1), 77–86.

Collins, P. H. (2015). Intersectionality's Definitional Dilemmas. *Annual Review of Sociology*, *41*, 1–20. doi:10.1146/annurev-soc-073014-112142

Daniel, R. M., Cousens, S. N., De Stavola, B. L., Kenward, M. G., & Sterne, J. A. C. (2013). Methods for dealing with time-dependent confounding. *Statistics in Medicine*, *32*(9), 1584–1618. doi:10.1002/sim.5686

De Stavola, B. L., Daniel, R. M., Ploubidis, G. B., & Micali, N. (2015). Mediation analysis with intermediate confounding: Structural equation modeling viewed through the causal inference lens. *American Journal of Epidemiology*, *181*(1), 64–80. doi:10.1093/aje/kwu239

Denier, N., Clouston, S. A. P., Richards, M., & Hofer, S. M. (2017). Retirement and cognition: A life course view.

Diez Roux, A. V. (2012). Conceptual Approaches to the Study of Health Disparities. *Annual Review of Public Health*, *33*(1), 41–58. doi:10.1146/annurev-publhealth-031811-124534

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Fisher, G. G., Hassan, H., Faul, J. D., Rodgers, W. L., & Weir, D. R. (2017). *Health and Retirement Study imputation of cognitive functioning measures: 1992–2014.* Ann Arbor, MI.

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

Fong, T. G., Ph, D., Fearing, M. a, Ph, D., Jones, R. N., Sc, D., et al. (2009). The Telephone

Interview for Cognitive Status: Creating a crosswalk with the Mini-Mental State Exam. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, *5*(6), 492–497. doi:10.1016/j.jalz.2009.02.007

Fonseca, R., Kapteyn, A., Lee, J., Zamarro, G., & Feeney, K. (2014). A Longitudinal Study of Well-Being of Older Europeans: Does Retirement Matter? *Journal of Population Ageing*, *7*(1), 21–41. doi:10.1007/s12062-014-9094-7

Ghisletta, P., Rabbitt, P., Lunn, M., & Lindenberger, U. (2012). Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence*, *40*(3), 260–268. doi:10.1016/j.intell.2012.02.008

Glymour, M. M., & Manly, J. (2008). Lifecourse social conditions and racial and ethnic patterns of cognitive aging. *Neuropsychol Rev*, *18*(3), 223–254. doi:10.1007/s11065-008-9064-z

González, H. M., Bowen, M. E., & Fisher, G. G. (2008). Memory Decline and Depressive Symptoms in a Nationally Representative Sample of Older Adults: The Health and Retirement Study (1998–2004). *Dementia & Geriatric Cognitive Disorders*, *25*(3), 266–271. doi:10.1159/000115976

Gow, A. J., Pattie, A., & Deary, I. J. (2016). Lifecourse Activity Participation From Early, Mid, and Later Adulthood as Determinants of Cognitive Aging: The Lothian Birth Cohort 1921. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *72*(1), gbw124. doi:10.1093/geronb/gbw124

Greenland, S., & Robins, J. M. (2009). Identifiability, exchangeability and confounding revisited. *Epidemiologic Perspectives & Innovations*, *6*(1), 4.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging. In F. I. Craik & S. Trehub (Eds.), *Aging and Cognitive*

Processes (pp. 237–263). New York: Plenum Press.

Hudomiet, P., Parker, A. M., & Rohwedder, S. (2018). Cognitive ability, personality, and pathways to retirement: An exploratory study. *Work, Aging and Retirement*, *4*(1), 52–66. doi:10.1093/workar/wax030

Jain, P., Danaei, G., Robins, J. M., Manson, J. E., Miguel, A., & Health, G. (2016). Smoking cessation and long-term weight gain in the Framingham Heart Study: an application of the parametric gformula for a continuous outcome. *Eur J Epidemiol*, *31*(12), 1223–1229. doi:10.1007/s10654-016-0200-4.Smoking

Jones, R. N., Manly, J., Glymour, M. M., Rentz, D. M., Jefferson, A. L., & Stern, Y. (2011). Conceptual and measurement challenges in research on Cognitive Reserve. *Journal of the International Neuropsychological Society*, *17*(04), 593–601. doi:10.1017/S1355617710001748

Kalleberg, A. L., Reskin, B. F., & Hudson, K. E. N. (2000). Bad Jobs in America : Standard and Nonstandard Employment Relations and Job Quality in the United States. *American Sociological Review*, *65*(2), 256–278.

Karlamangla, A. S., Miller-Martinez, D., Aneshensel, C. S., Seeman, T. E., Wight, R. G., & Chodosh, J. (2009). Trajectories of cognitive function in late life in the United States: Demographic and socioeconomic predictors. *American Journal of Epidemiology*, *170*(3), 331–342. doi:10.1093/aje/kwp154

Keil, A. P., Edwards, J. K., Richardson, D. B., Naimi, A. I., & Cole, S. R. (2014). The parametric g-formula for time-to-event data: Intuition and a worked example. *Epidemiology*, *25*(6), 889–897. doi:10.1097/EDE.0000000000000160

Kuiper, J. S., Zuidersma, M., Oude Voshaar, R. C., Zuidema, S. U., van den Heuvel, E. R., Stolk,

R. P., & Smidt, N. (2015). Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews*, *22*(June), 39–57. doi:10.1016/j.arr.2015.04.006

Lin, S. H., Young, J. G., Logan, R., & VanderWeele, T. J. (2017). Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders. *Statistics in Medicine*, *36*(26), 4153–4166. doi:10.1002/sim.7426

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

Mehta, N., & Preston, S. (2016). Are major behavioral and sociodemographic risk factors for mortality additive or multiplicative in their effects? *Social Science and Medicine*, *154*, 93–99. doi:10.1016/j.socscimed.2016.02.009

Meng, A., Nexø, M. A., & Borg, V. (2017). The impact of retirement on age related cognitive decline - A systematic review. *BMC Geriatrics*, *17*(1), 1–10. doi:10.1186/s12877-017-0556-7

Moffitt, R. (2005). Remarks on the analysis of causal relationships in population research. *Demography*, *42*(1), 91–108. doi:10.1353/dem.2005.0006

Mosca, I., & Wright, R. E. (2018). Effect of Retirement on Cognition: Evidence From the Irish Marriage Bar. *Demography*, *55*(4), 1317–1341. doi:10.1007/s13524-018-0682-7)

Petersen, M. L., Porter, K. E., Gruber, S., Wang, Y., & Van Der Laan, M. J. (2012). Diagnosing and responding to violations in the positivity assumption. *Statistical methods in medical research*, *21*(1), 31–54.

Potter, G. G., Helms, M. J., & Plassman, B. L. (2008). Associations of job demands and intelligence with cognitive performance among men in late life. *Neurology*, *70*(19 Part 2),

1803–1808. doi:10.1212/01.wnl.0000295506.58497.7e

RAND Center for the Study of Aging. (2017). RAND HRS Data, Version P. Santa Monica, CA.

Reed, B. R., Dowling, M., Tomaszewski Farias, S., Sonnen, J., Strauss, M., Schneider, J. A., et al. (2011). Cognitive activities during adulthood are more important than education in building reserve. *Journal of the International Neuropsychological Society*, *17*(04), 615–624. doi:10.1017/S1355617711000014

Rehkopf, D. H., Glymour, M. M., & Osypuk, T. L. (2016). The consistency assumption for causal inference in social epidemiology: when a rose is not a rose. *Current epidemiology reports*, *3*(1), 63–71.

Roberts, B. A., Fuhrer, R., Marmot, M., & Richards, M. (2011). Does retirement influence cognitive performance? The Whitehall II Study. *Journal of Epidemiology and Community Health*, *65*(11), 958–963. doi:10.1136/jech.2010.111849

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, *7*(9–12), 1393–1512.

Rohwedder, S., & Willis, R. J. (2010). Mental Retirement. *Journal of Economic Perspectives*, *24*(1), 119–138. doi:10.1257/jep.24.1.119

Schoeni, R. F., Freedman, V. A., & Langa, K. M. (2018). Introduction to a Supplement on Population Level Trends in Dementia: Causes, Disparities, and Projections. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *73*(August), S1–S9. doi:10.1093/geronb/gby007

Sharp, B. K. (2017). Stress as experienced by people with dementia: An interpretative phenomenological analysis. *Dementia*, 147130121771387. doi:10.1177/1471301217713877

Staudinger, U. M., Finkelstein, R., Calvo, E., & Sivaramakrishnan, K. (2016). A global view on the effects of work on health in later life. *Gerontologist*, *56*, S281–S292. doi:10.1093/geront/gnw032

University of Michigan. (2017). Health and Retirement Study Public Use Dataset. Ann Arbor, MI.

Vanderweele, T. J., & Tchetgen Tchetgen, E. (2017). Mediation Analysis with Time-Varying Exposures and Mediators. *Harvard Univerisity Biostatistics Working Paper Series*, 1–22.

VanderWeele, T. J., & Tchetgen Tchetgen, E. J. (2017). Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *79*(3), 917–938. doi:10.1111/rssb.12194

VanderWeele, T. J., Vansteelandt, S., & Robins, J. M. (2014). Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology (Cambridge, Mass.)*, *25*(2), 300.

Virtanen, M., Ferrie, J. E., Batty, G. D., Elovainio, M., Jokela, M., Vahtera, J., et al. (2015). Socioeconomic and psychosocial adversity in midlife and depressive symptoms post retirement: A 21-year follow-up of the whitehall II study. *American Journal of Geriatric Psychiatry*, *23*(1), 99-109.e1. doi:10.1016/j.jagp.2014.04.001

Wang, A., & Arah, O. A. (2015). G-computation demonstration in causal mediation analysis. *European Journal of Epidemiology*, *30*(10), 1119–1127. doi:10.1007/s10654-015-0100-z

Appendix I

Table A-1 Age-specific population-averaged effect and average treatment effect on the treated of postponing retirement on cognitive function with 95% confidence intervals

	Population-Averaged Effect			**Average Treatment on the Treated**		
Age	Effect	95% CI		Effect	95% CI	
57	0.01	0.00	0.02	0.10	0.03	0.16
58	0.03	0.01	0.04	0.17	0.08	0.26
59	0.05	0.03	0.08	0.22	0.12	0.32
60	0.08	0.04	0.12	0.25	0.13	0.36
61	0.11	0.07	0.16	0.28	0.17	0.40
62	0.15	0.09	0.21	0.30	0.19	0.42
63	0.18	0.12	0.26	0.32	0.21	0.45
64	0.22	0.14	0.31	0.34	0.22	0.47
65	0.26	0.17	0.36	0.36	0.23	0.49
66	0.30	0.18	0.41	0.38	0.23	0.52
67	0.34	0.20	0.47	0.43	0.26	0.60
68	0.30	0.18	0.41	0.38	0.23	0.52
69	0.23	0.15	0.32	0.30	0.20	0.42
70	0.18	0.12	0.24	0.23	0.15	0.31
71	0.13	0.09	0.18	0.17	0.12	0.24
72	0.10	0.06	0.13	0.13	0.08	0.18
73	0.07	0.04	0.10	0.09	0.06	0.14
74	0.05	0.03	0.07	0.07	0.04	0.10
75	0.04	0.02	0.06	0.05	0.03	0.08

Table A-2 Age-specific population-averaged effect (PAE) and average treatment effect on the treated (ATT) of postponing retirement until at least age 67 on cognitive function by gender (Panel A), educational attainment (Panel B), and occupational attainment (Panel C) with 95% confidence intervals

Panel A. By gender

	Population-Averaged Effect						Average Treatment on the Treated					
Age	Men	95% CI		Women	95% CI		Men	95% CI		Women	95% CI	
57	0.01	0.00	0.01	0.01	0.00	0.02	0.10	0.02	0.17	0.10	0.01	0.19
58	0.02	0.01	0.04	0.03	0.01	0.05	0.17	0.06	0.26	0.18	0.07	0.28
59	0.05	0.02	0.08	0.06	0.03	0.08	0.22	0.10	0.35	0.22	0.10	0.33
60	0.07	0.03	0.11	0.09	0.05	0.13	0.25	0.11	0.39	0.25	0.14	0.36
61	0.10	0.06	0.16	0.12	0.07	0.18	0.28	0.16	0.41	0.28	0.16	0.40
62	0.13	0.08	0.20	0.16	0.10	0.23	0.30	0.16	0.43	0.30	0.18	0.43
63	0.17	0.10	0.24	0.20	0.12	0.28	0.32	0.18	0.45	0.32	0.20	0.45
64	0.20	0.12	0.29	0.23	0.14	0.33	0.33	0.19	0.48	0.34	0.21	0.48
65	0.23	0.12	0.35	0.27	0.17	0.38	0.35	0.18	0.50	0.36	0.22	0.50
66	0.27	0.15	0.40	0.32	0.19	0.44	0.36	0.20	0.52	0.39	0.24	0.54
67	0.31	0.16	0.45	0.36	0.22	0.52	0.42	0.22	0.59	0.44	0.27	0.63
68	0.27	0.15	0.38	0.31	0.20	0.45	0.37	0.20	0.52	0.39	0.24	0.55
69	0.22	0.12	0.30	0.25	0.16	0.35	0.30	0.17	0.42	0.31	0.20	0.44
70	0.16	0.10	0.22	0.19	0.13	0.26	0.23	0.13	0.31	0.24	0.16	0.33
71	0.12	0.07	0.17	0.14	0.09	0.19	0.17	0.10	0.24	0.18	0.12	0.24
72	0.09	0.05	0.12	0.10	0.06	0.15	0.12	0.07	0.18	0.13	0.08	0.19
73	0.06	0.03	0.10	0.07	0.04	0.11	0.09	0.05	0.14	0.10	0.05	0.14
74	0.05	0.02	0.07	0.06	0.03	0.08	0.07	0.03	0.11	0.07	0.04	0.11
75	0.03	0.01	0.05	0.04	0.02	0.07	0.05	0.02	0.08	0.06	0.03	0.09

Panel B. By educational attainment:

	Population-Averaged Effect									Average Treatment on the Treated								
Age	<HS or GED	95% CI		HS	95% CI		Some	95% CI		<HS or GED	95% CI		HS	95% CI		Some	95% CI	
57	0.01	0.00	0.02	0.01	0.00	0.02	0.01	0.00	0.02	0.07	-0.02	0.17	0.11	0.03	0.18	0.12	0.01	0.22
58	0.02	0.00	0.05	0.03	0.01	0.05	0.03	0.01	0.04	0.11	0.00	0.25	0.18	0.07	0.29	0.22	0.09	0.34
59	0.04	0.00	0.08	0.06	0.03	0.09	0.05	0.02	0.08	0.14	-0.01	0.27	0.24	0.13	0.34	0.27	0.13	0.40
60	0.06	0.01	0.11	0.09	0.06	0.13	0.08	0.04	0.11	0.17	0.02	0.30	0.27	0.16	0.38	0.30	0.16	0.43
61	0.09	0.02	0.15	0.13	0.08	0.18	0.11	0.06	0.16	0.19	0.04	0.33	0.30	0.18	0.41	0.33	0.19	0.47
62	0.11	0.03	0.18	0.17	0.11	0.23	0.15	0.09	0.21	0.20	0.05	0.33	0.32	0.21	0.45	0.36	0.22	0.53
63	0.14	0.04	0.22	0.21	0.13	0.28	0.19	0.12	0.27	0.22	0.06	0.35	0.34	0.22	0.46	0.38	0.24	0.53
64	0.16	0.05	0.26	0.24	0.15	0.33	0.23	0.14	0.32	0.23	0.07	0.37	0.36	0.22	0.50	0.39	0.25	0.55
65	0.19	0.07	0.30	0.28	0.17	0.39	0.27	0.18	0.38	0.25	0.09	0.40	0.38	0.23	0.53	0.42	0.28	0.58
66	0.22	0.08	0.35	0.33	0.19	0.46	0.31	0.20	0.44	0.27	0.10	0.42	0.40	0.23	0.57	0.43	0.28	0.61
67	0.25	0.07	0.39	0.37	0.21	0.51	0.35	0.23	0.51	0.31	0.09	0.48	0.46	0.26	0.64	0.50	0.32	0.71
68	0.23	0.07	0.36	0.32	0.18	0.43	0.31	0.20	0.43	0.28	0.09	0.45	0.40	0.23	0.54	0.44	0.28	0.61
69	0.19	0.07	0.30	0.25	0.15	0.34	0.25	0.15	0.35	0.23	0.09	0.38	0.31	0.19	0.43	0.36	0.22	0.50
70	0.15	0.06	0.22	0.18	0.11	0.25	0.19	0.12	0.27	0.18	0.07	0.28	0.23	0.15	0.32	0.28	0.17	0.39
71	0.11	0.04	0.17	0.13	0.08	0.18	0.15	0.09	0.21	0.14	0.05	0.21	0.17	0.10	0.23	0.21	0.13	0.30
72	0.08	0.03	0.12	0.09	0.06	0.13	0.11	0.06	0.16	0.11	0.04	0.16	0.12	0.07	0.16	0.16	0.09	0.23
73	0.06	0.02	0.10	0.07	0.03	0.10	0.08	0.04	0.13	0.08	0.03	0.13	0.09	0.04	0.13	0.12	0.06	0.19
74	0.05	0.01	0.08	0.05	0.02	0.07	0.06	0.03	0.10	0.06	0.02	0.10	0.06	0.03	0.09	0.09	0.05	0.15
75	0.04	0.00	0.07	0.04	0.01	0.06	0.05	0.02	0.08	0.05	0.01	0.09	0.05	0.02	0.08	0.07	0.03	0.11

Panel C. By occupational attainment

	Population-Averaged Effect						Average Treatment on the Treated					
Age	Professional	95% CI		Non-Professional	95% CI		Professional	95% CI		Non-Professional	95% CI	
57	0.01	0.00	0.01	0.01	0.00	0.02	0.09	-0.01	0.19	0.11	0.04	0.17
58	0.02	0.01	0.04	0.03	0.01	0.05	0.16	0.04	0.27	0.18	0.06	0.27
59	0.04	0.01	0.06	0.06	0.03	0.09	0.20	0.06	0.31	0.23	0.13	0.33
60	0.06	0.03	0.10	0.09	0.05	0.13	0.23	0.10	0.35	0.26	0.15	0.38
61	0.09	0.04	0.13	0.12	0.08	0.18	0.25	0.12	0.38	0.29	0.18	0.42
62	0.11	0.06	0.18	0.16	0.10	0.23	0.27	0.13	0.41	0.31	0.20	0.43
63	0.14	0.07	0.22	0.20	0.14	0.28	0.28	0.14	0.43	0.33	0.22	0.46
64	0.17	0.09	0.26	0.24	0.15	0.34	0.29	0.15	0.45	0.35	0.23	0.49
65	0.20	0.09	0.32	0.28	0.18	0.40	0.31	0.14	0.48	0.37	0.24	0.54
66	0.24	0.11	0.36	0.32	0.20	0.46	0.32	0.15	0.49	0.40	0.25	0.57
67	0.27	0.12	0.41	0.37	0.21	0.52	0.37	0.16	0.57	0.45	0.27	0.64
68	0.24	0.11	0.36	0.32	0.19	0.45	0.34	0.16	0.51	0.40	0.23	0.56
69	0.20	0.10	0.30	0.25	0.16	0.35	0.28	0.13	0.42	0.32	0.20	0.44
70	0.16	0.08	0.23	0.19	0.12	0.26	0.22	0.11	0.33	0.24	0.16	0.33
71	0.12	0.07	0.17	0.14	0.09	0.19	0.17	0.09	0.25	0.17	0.12	0.24
72	0.09	0.05	0.13	0.10	0.06	0.14	0.13	0.07	0.19	0.13	0.08	0.19
73	0.07	0.03	0.10	0.07	0.04	0.11	0.09	0.05	0.15	0.09	0.05	0.14
74	0.05	0.02	0.08	0.05	0.03	0.08	0.07	0.04	0.12	0.07	0.04	0.10
75	0.04	0.01	0.07	0.04	0.02	0.06	0.06	0.02	0.10	0.05	0.03	0.08

Appendix II

In sensitivity analyses, we also examine scenarios whereby instead of "allowing" respondents in the intervention scenario to retire at age 67 following a natural course scenario (meaning they exit the labor force in phases), we "force" everyone to retire at age 67. We present figures below displaying both the population-averaged effects (PAE) and average treatment effect on the treated (ATT) by gender, education, and occupation.

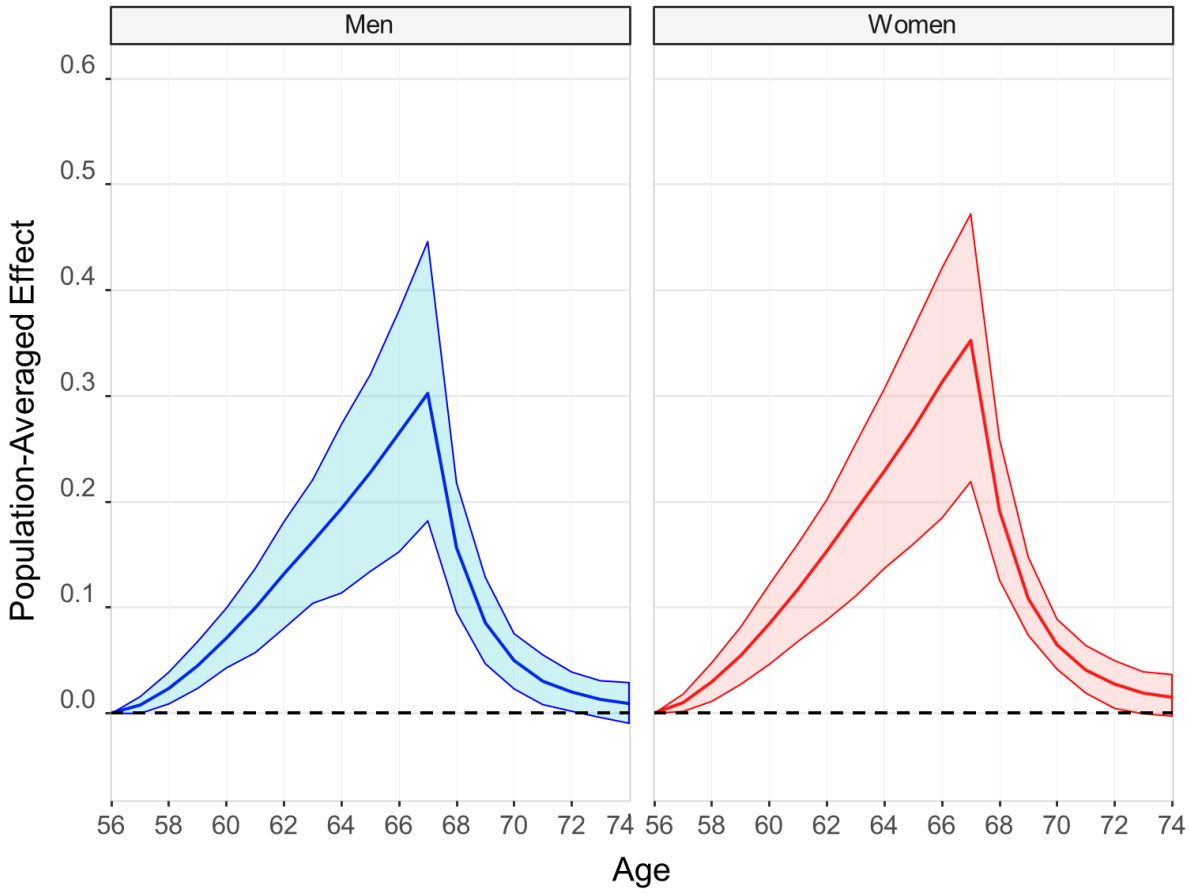

Figure A-1 Population-averaged effect by gender—the difference in cognitive functioning between the natural course scenario and the intervention, whereby retirement is postponed until age 67, and then all are "forced" to retire at age 67.

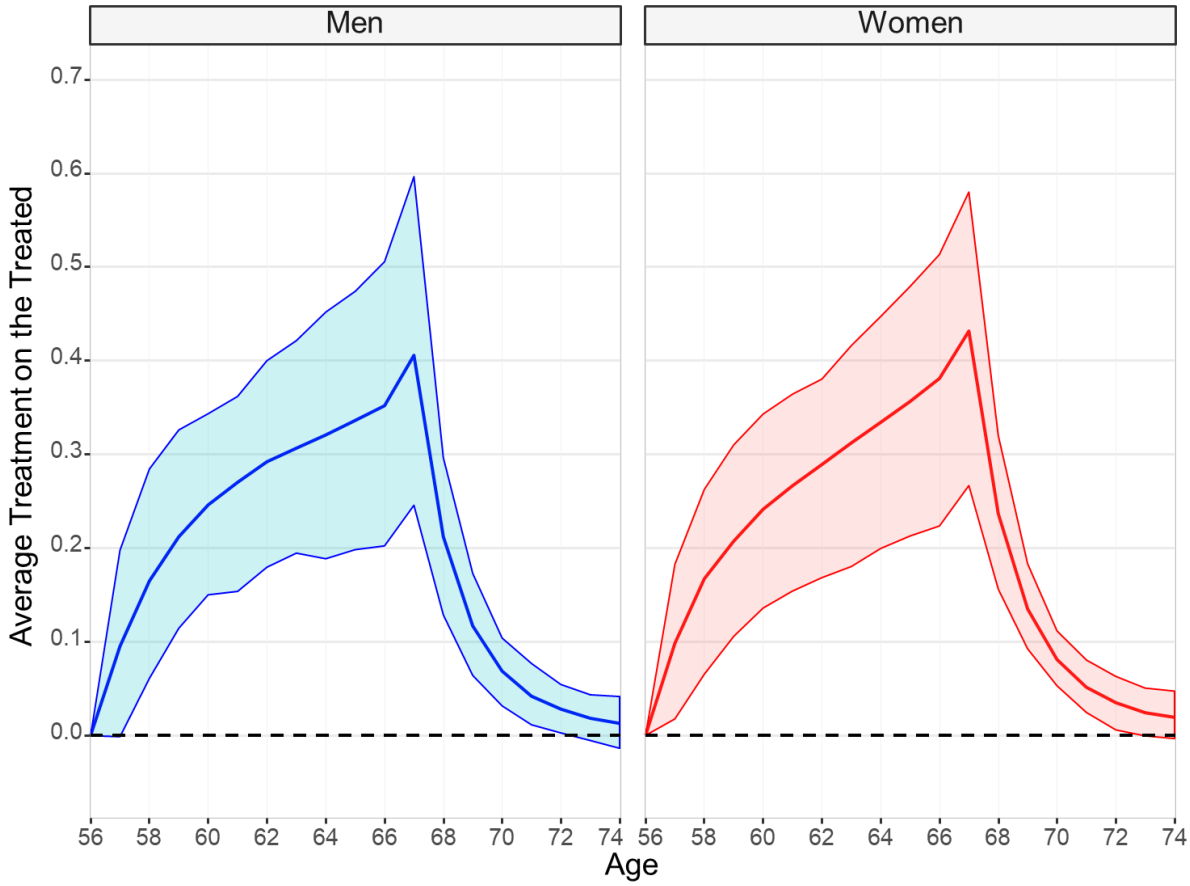

Figure A-2 Average treatment effect on the treated by gender— the difference in cognitive function score between the natural course and intervention scenarios where we postpone retirement until age 67 for those who retired prior to age 67 and then "force" them to retire at age 67.

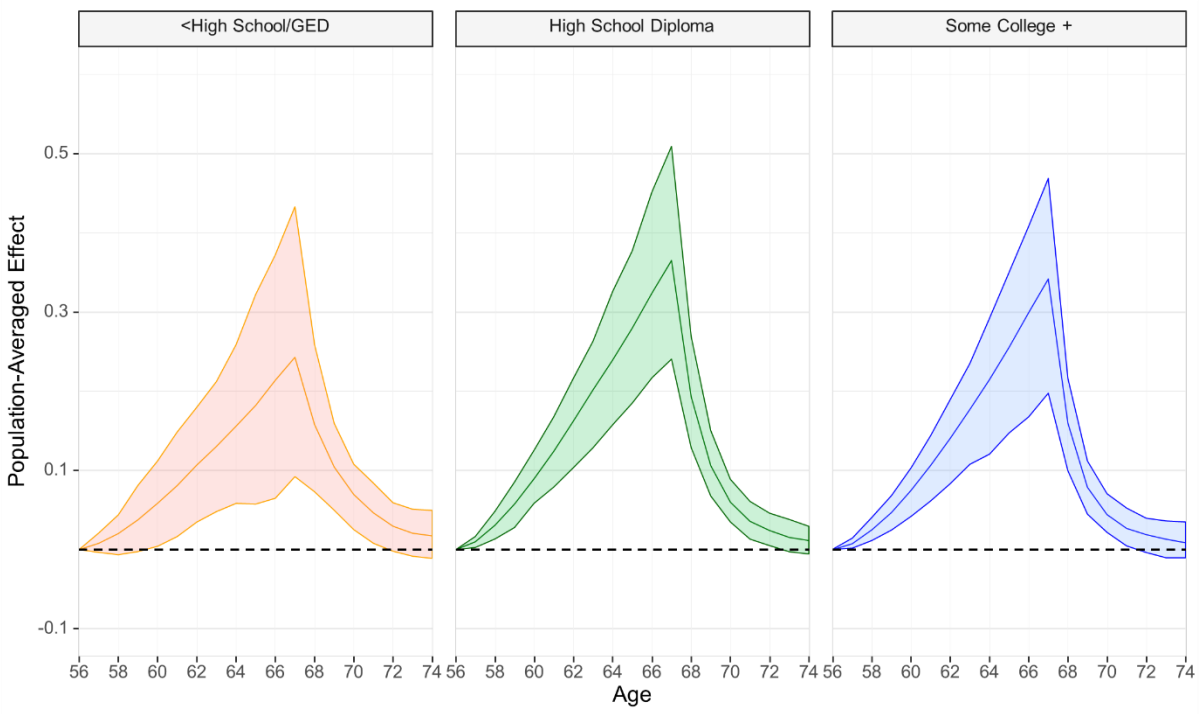

Figure A-3 Population-averaged effect by educational attainment—the difference in cognitive functioning between the natural course scenario and the intervention, whereby retirement is postponed until age 67, and then all are "forced" to retire at age 67.

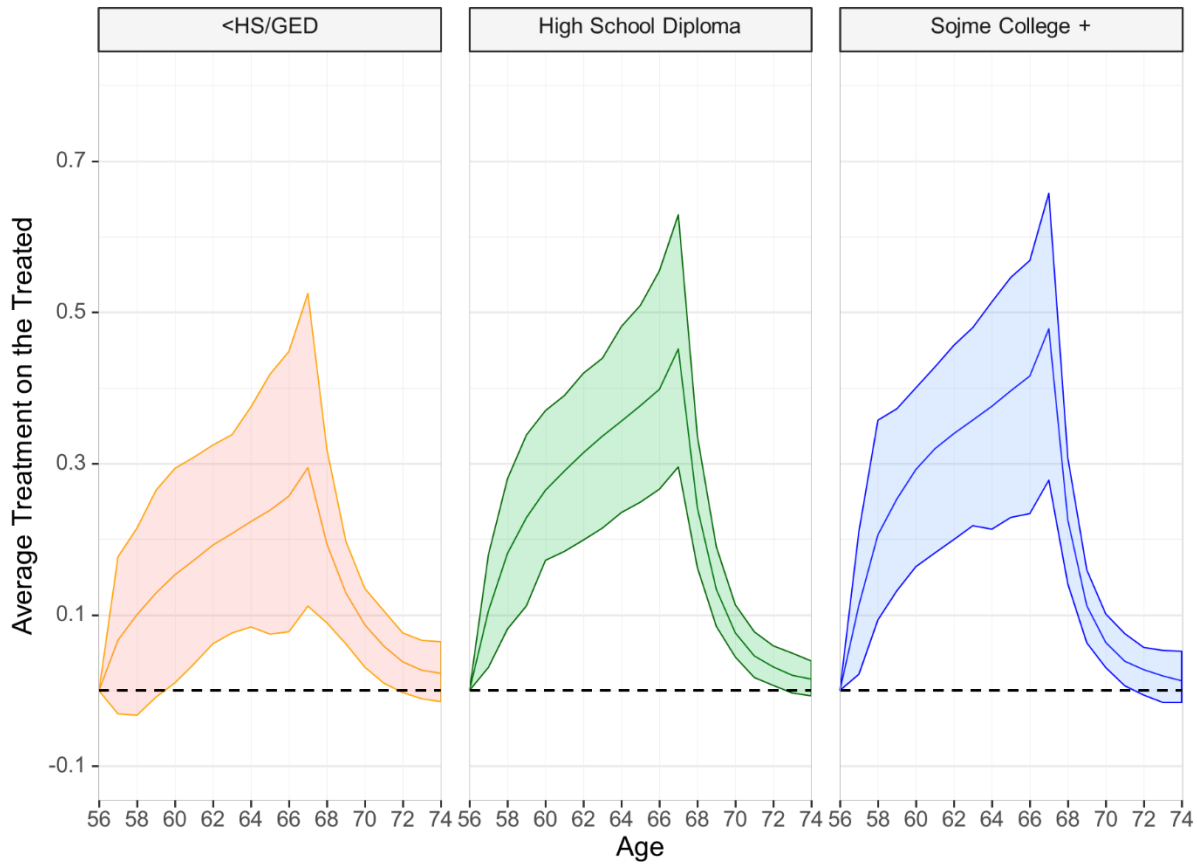

Figure A-4 Average treatment effect on the treated by educational attainment— the difference in cognitive function score between the natural course and intervention scenarios where we postpone retirement until age 67 for those who retired prior to age 67 and then "force" them to retire at age 67.

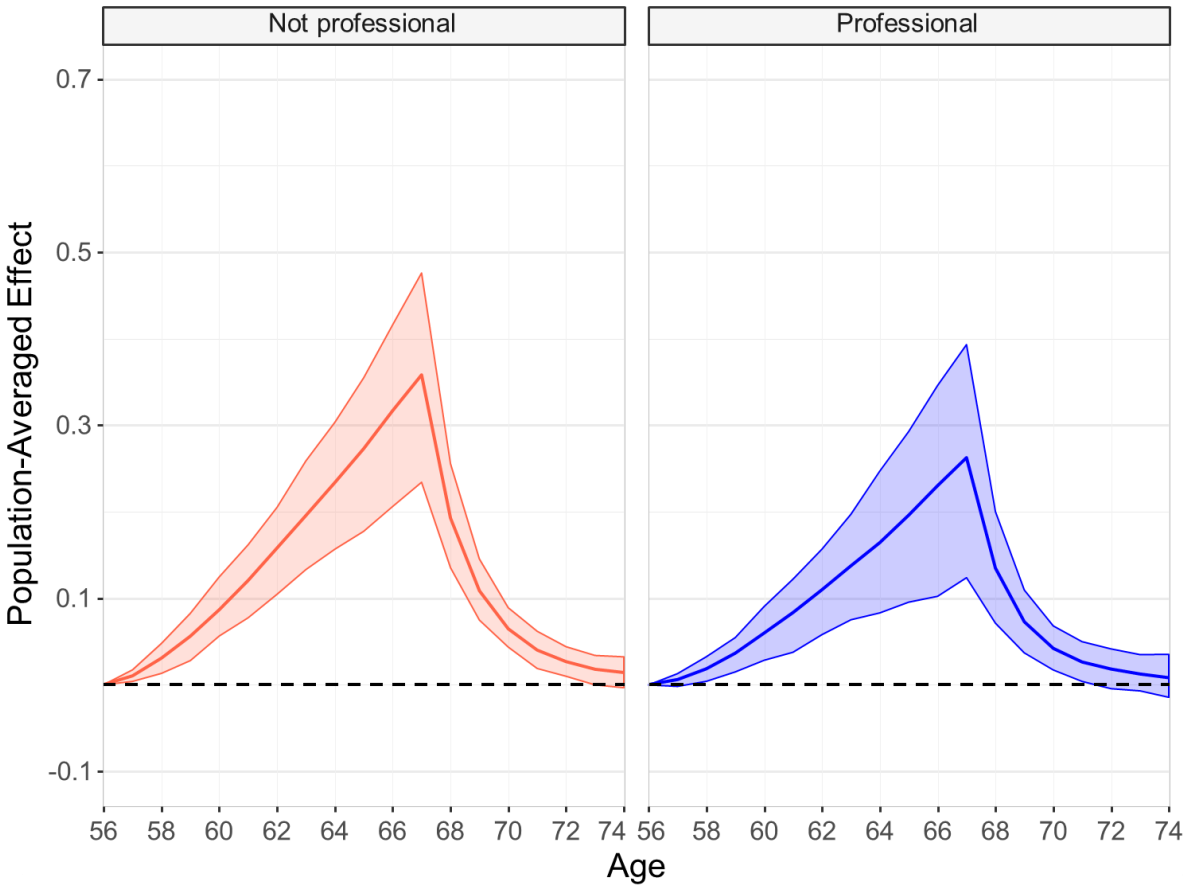

Figure A-5 Population-averaged effect by occupational attainment—the difference in cognitive functioning between the natural course scenario and the intervention, whereby retirement is postponed until age 67, and then all are "forced" to retire at age 67.

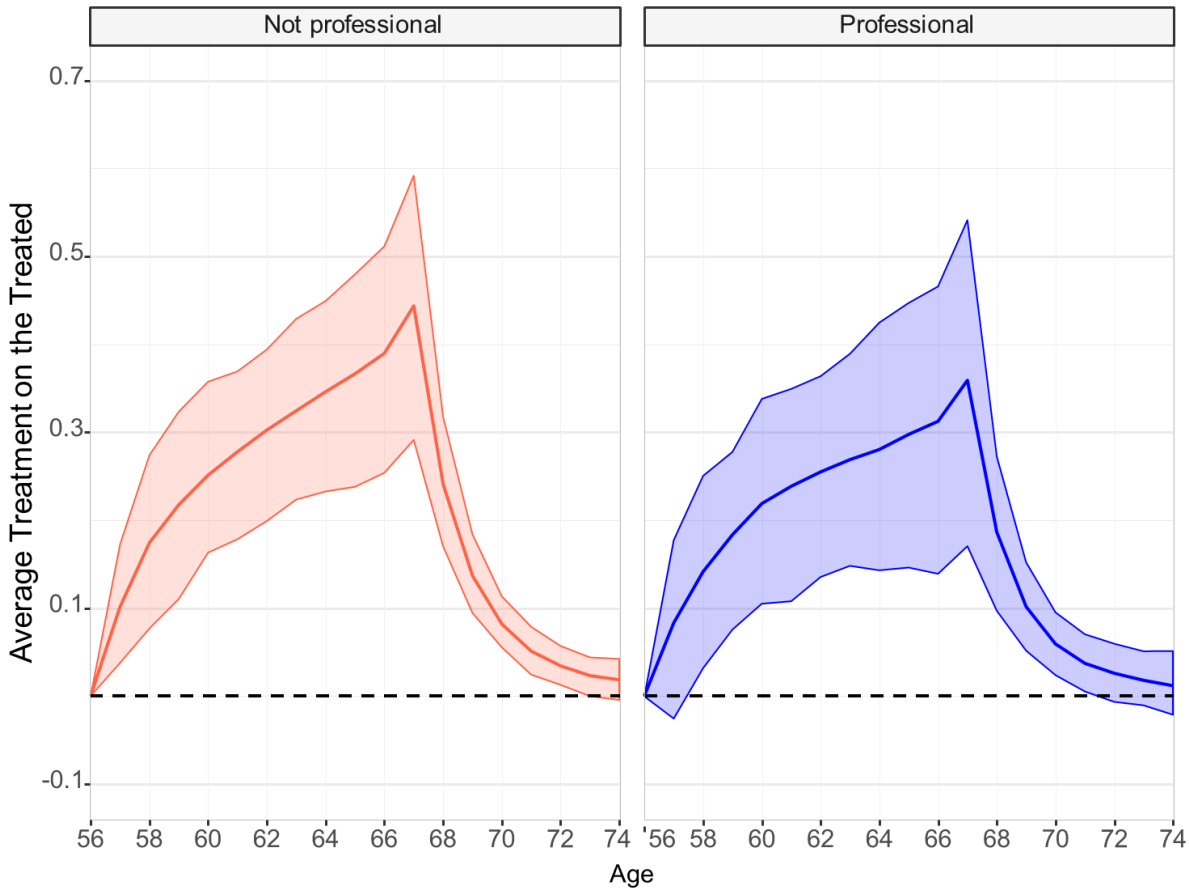

Figure A-6 Average treatment effect on the treated by occupational attainment— the difference in cognitive function score between the natural course and intervention scenarios where we postpone retirement until age 67 for those who retired prior to age 67 and then "force" them to retire at age 67.

Appendix III

Mediation definitions

Hypothesis 3: Depressive symptoms and comorbidities act as mechanisms, mediating the effect of postponing retirement on later-life cognitive function.

To test this hypothesis, we perform mediation analysis. First, we compare the natural course scenario (an approximation of the empirical data) with the intervention scenario (where retirement is postponed to age 67 for all). In counterfactual notation (Wang and Arah 2015), this is shown as follows:

$$E[Y_R] = E[Y_{R,D_R,C_R}]$$

Where E is the mathematical expectation, Y represents to the outcome variable cognition, R refers to retirement status, D to depression score, and C to comorbidity score. Taken together, this represents the average outcome Y when R is kept as observed, and D and C take on the value they have when R is as observed.

In counterfactual notation, the intervention scenario is shown as follows:

$$E[Y_{R^*}] = E[Y_{R^*,D_{R^*},C_{R^*}}]$$

Where the star represents an intervention on R. Since D and C have R^* in their subscript, this represents the value they would take when there has been an intervention on R.

Subtracting the average cognition score in the natural course scenario from the average cognition score in the intervention scenario, we find the Total Effect (TE). Whether this total effect is a population-averaged effect (PAE) or an Average Treatment Effect for the Treated (ATT) is dependent on the subpopulation for which the average cognitive function score is calculated. Using counterfactual notation, the TE is shown as follows:

Total Effect = TE : $$E\left[Y_{R^*,D_{R^*},C_{R^*}}\right] - E\left[Y_{R,D_R,C_R}\right]$$

To perform mediation analysis, we simulate two additional scenarios. In both scenarios, the intervention (postponement of retirement to age 67) is performed, but a set of mediating variables is kept at the natural course levels, i.e., the intervention will not affect these mediating variables, hence the intervention will not act through the mediating variables onto the outcome variable (cognition). The intervention will only act on the outcome variable directly and through mediating variables that were not kept at the natural course levels.

In one scenario, depression score (CESD) is kept at natural course levels. In counterfactual notation, this is shown as:

$$E[Y_{R^*,D_R,C_{R^*}}]$$

In another scenario, comorbidity score is kept at natural course levels. In counterfactual notation:

$$E[Y_{R^*,D_{R^*},C_R}]$$

Using this notation, the direct effect of the intervention (not via CESD) is calculated as:

Direct effect not via CESD: $E\left[Y_{R^*,D_R,C_{R^*}}\right] - E[Y_{R,D_R,C_R}]$

And the direct effect of the intervention (not via comorbidity) is calculated as:

Direct effect not via comorbidity: $E[Y_{R^*,D_{R^*},C_R}] - E[Y_{R,D_R,C_R}]$

The indirect effect of the intervention via CESD is the total effect minus the direct effect not via CESD, and hence is calculated as:

Indirect effect via CESD: $\left\{E\left[Y_{R^*,D_{R^*},C_{R^*}}\right] - E\left[Y_{R,D_R,C_R}\right]\right\} - \left\{E\left[Y_{R^*,D_R,C_{R^*}}\right] - E[Y_{R,D_R,C_R}]\right\}$

And the indirect effect of the intervention via comorbidity is similarly calculated:

Indirect effect via comorbidity: $\left\{E\left[Y_{R^*,D_{R^*},C_{R^*}}\right] - E\left[Y_{R,D_R,C_R}\right]\right\} - \left\{E[Y_{R^*,D_{R^*},C_R}] - E[Y_{R,D_R,C_R}]\right\}$

In Table A3, the numbers represent the direct effect relative to the total effect, i.e., the proportion of the effect that does not operate via the mediators in question. An effect of 1 means that the total effect is entirely a direct effect. An effect lower than 1, e.g., 0.90, means that 90% of the total effect is direct, and 10% of the total effect is via the mediators (and the effect via the mediators is in the same direction as the direct effect). In other words, 10% of the total effect is because of how postponed retirement affects depression (or comorbidities), which in turn affects cognitive function. A direct effect of 1.10 means that (1/1.10 = 0.909, 1 – 0.909 = 0.091) 9.1% of the total effect is via mediators, and the effect of the mediators is in the opposite direction of direct effect. For example, if the effect is greater than 1, postponing retirement is insulative against cognitive decline, but postponed retirement increases depressive symptoms or

comorbidities, which are in turn associated with cognitive decline, and therefore the overall protective effect of postponed retirement is lessened. Table A3, however, shows a non-significant mediated effect at almost all ages for both depressive symptoms and comorbidities.

Table A-3 The direct effect, operating either not via depressive symptoms or not via comorbidities

	Not via Depressive symptoms			Not via Comorbidities		
Age	**Direct effect**	**95% CI**		**Direct effect**	**95% CI**	
57	0.98	0.55	1.60	1.02	0.57	1.67
58	1.09	0.86	1.38	1.04	0.79	1.31
59	1.12	0.99	1.29	1.06	0.91	1.21
60	1.11	0.99	1.25	1.05	0.95	1.16
61	1.09	1.01	1.19	1.04	0.95	1.15
62	1.06	1.00	1.14	1.04	0.96	1.11
63	1.04	0.97	1.11	1.03	0.96	1.10
64	1.03	0.97	1.09	1.03	0.98	1.10
65	1.01	0.96	1.06	1.02	0.97	1.07
66	0.98	0.94	1.03	1.01	0.97	1.05
67	0.96	0.91	1.01	1.00	0.96	1.05
68	0.94	0.88	1.00	1.00	0.94	1.05
69	0.92	0.84	0.99	0.99	0.91	1.08
70	0.91	0.82	1.00	0.97	0.87	1.09
71	0.90	0.78	1.01	0.94	0.79	1.10
72	0.89	0.74	1.02	0.89	0.71	1.10
73	0.92	0.71	1.19	0.84	0.48	1.23
74	0.92	0.63	1.25	0.73	0.29	1.27