

Effects of working part-time and full-time on physical and mental health in old age in Europe

Tunga Kantarcı Ingo Kolodziej

Tilburg University and Netspar

RWI - Leibniz Institute for Economic Research and Ruhr-University Bochum

Utrecht, 14 October 2016

Introduction

- A growing body of literature is analyzing if **retirement** has an effect on mental and physical health: Mazzonna and Peracchi, 2010; Rohwedder and Willis, 2010; Bonsang et al., 2011; Coe and Zamarro, 2011; Insler, 2014; Eibich, 2015; Hallberg et al., 2015; Godard, 2016; Kämpfen and Maurer, 2016
- Some find, e.g., that those who are **retired** have better overall health, or score lower in cognition tests than those who are **working any number of hours**
- These studies estimate models using IV to circumvent the problem that people with health problems may select themselves into retirement

Introduction

- If **retirement** has an effect on health, should working **some number of hours** necessarily have the opposite effect?
- What is the amount of work hours that preserves or deteriorates health?
- Does health respond to work hours in a linear fashion?
- To answer these questions we study the effects of working part-time and full-time on physical and mental health in old age

Introduction

- Several studies analyzed if working part-time in old age has an effect on health in the US
 - Dave et al. (2008) find that part-time workers or retirees have worse physical health than fully retired people
 - Liu et al. (2009) find that part-time workers or retirees have fewer major diseases and functional limitations than fully retired people
 - Neuman (2008) shows that a reduction in the number of work hours from full-time preserves general and physical health

Introduction

- These studies take different approaches to control for selection
 - Dave et al. consider the effect of working for those who did not have a health problem in the survey years prior to retirement
But this assumes that changes in health status in between the biennial survey years or in the current survey year has not affected work decisions in the current survey year. Besides, odds of having a health problem in future may be higher for those who never had a health problem before
 - Liu et al. consider the effect of current work status on future health status
But this assumes that expectations for future health status do not affect current work decisions
 - Neuman uses retirement eligibility ages as instruments for number of work hours
But Neuman considers working less than 3 days a week as full retirement, but this allows making claims on reducing hours, not on part-time working

Empirical approach

- We could estimate the effects of working part-time and full-time on health by OLS in the equation:

$$Y_{it} = \alpha + f(S_{it}) + D_{it}^j \beta + u_{it}$$

Y_{it} is a measure of health. S_{it} is age of the individual. $f(S_{it})$ is a continuous function of age that controls for age related changes in the health status. D_{it}^j is a vector of dummy variables for part-time and full-time work status denoted by j . β is the parameter of interest which measures the response of health to working part-time and full-time

Empirical approach - Unobserved heterogeneity

- D_{it}^j can be endogenous because of time-invariant idiosyncratic characteristics that are correlated with the health outcome as well as the retirement behavior. We allow for fixed effects to control for unobserved heterogeneity:

$$Y_{it} = \alpha + f(S_{it}) + D_{it}^j \beta + \mu_i + u_{it}$$

μ_i is a time-invariant individual specific unobserved variable and it is potentially correlated with D_{it}^j and with S_{it} .

Empirical approach - Selection

- D_{it}^j can still be endogenous due to selection. E.g., Gannon and Roberts (2011) and Bound et al. (1999) have shown that elderly people who have a health problem reduce their work hours or stop working all together. We take an IV approach to circumvent selection. The estimation consists of two stages. In the first stage we estimate

$$D_{it}^j = f^j(S_{it}) + I(S_{it} \geq \bar{S})\gamma^j + \eta_i^j + \epsilon_{it}^j$$

$f^j(S_{it})$ is a continuous function of age. \bar{S} is the vector of early and normal retirement eligibility ages for social security benefits, and the vector $I(S_{it} \geq \bar{S})$ indicates whether the individual is at least as old as each of these eligibility ages. γ^j measures the discontinuities in the probabilities of working part-time or full-time at the eligibility ages \bar{S} . η_i^j are fixed effects potentially correlated with age

Empirical approach - Selection

- To be valid instruments, retirement eligibility ages are required to be relevant predictors of part-time and full-time work decisions, and exogenous to the health status of the individual
 - It is well documented that retirement ages are strong predictors of retirement decisions (Hurd, 1990; Mitchell, 1999).
 - We also use the retirement ages of the spouse as instruments since several studies provide empirical evidence that couples coordinate their retirement timing (Blau, 1998; Gustman and Steinmeier, 2000, 2004).
 - Health status is not likely to change at the institutionally set eligibility ages of the individual or the spouse
- Retirement ages are typical instruments in the subject literature (Charles, 2004; Rohwedder and Wills, 2009; Coe and Zamarro, 2011, etc.)

Empirical approach - Selection

- In the second stage we estimate

$$\tilde{Y}_{it} = \tilde{f}(S_{it}) + \hat{D}_{it}^j \beta + \tilde{v}_{it}$$

\hat{D}_{it}^j are the within group transformed part-time and full-time work probabilities predicted in the first stage regression

- That is, we estimate a fixed effects model with instrumental variables

Data - Measures of health

- Self-perceived health
 - *“Would you say your health is very good, good, fair, bad, and very bad?”*
 - Self-perceived health may bias the effect of work hours on health because respondents may report an inferior health status to justify their work status, or individuals may differ in their response scales as they give subjective judgments of their own health status and cause measurement error

Data - Measures of health

- Health index
 - Following Coe and Zamarro (2011), we create a *health stock* variable by predicting self-perceived health by objective physical and mental health measures that is less prone to reporting bias:

$$H_{it} = \alpha + L_{it}\beta + \nu_i + \mu_{it}$$

H_{it} is the self-perceived health status. ν_i controls for unobserved heterogeneity. L_{it} is a vector of objective measures of health which include the number of limitations in the activities and instrumental activities of daily living (ADL, IADL), total number of chronic diseases, a summary index of mobility, any overnight hospital stay within the last two years, body mass index, scores of the word recall and subtraction tests, a summary index of depression (EURO-D)

Data - Measures of health

FE model explaining self-perceived health

	Self-perceived health	
	Coefficient	p-value
Number of ADL limitations	0.054	0.000
Number of IADL limitations	-0.013	0.478
Number of mobility limitations	0.112	0.000
Number of difficulties in muscle use	0.097	0.000
Number of chronic diseases	0.121	0.000
Hospital stay	0.198	0.000
Word recall test	-0.002	0.071
Numeracy	0.012	0.005
Depression	0.062	0.000
Fluency	-0.004	0.000
Constant	2.525	0.000
F-test for overall significance		0.000
N obs.	80119	
N ind.	46766	
Model: Linear model with FE		

Notes: Self-perceived health: 1 (very good), ..., 5 (very bad). SE robust to heteroskedasticity and clustering on panel respondents.

Data - Measures of health

- Body mass index.
 - $BMI = Weight / Height^2$
- Word recall
 - Respondents are presented with a list of 10 words to memorize. They are then asked immediately to recall as many words as possible from the list. After asking other questions, they are asked for a second time to recall as many words as possible from the same list
 - Each immediate or delayed recall of a word is counted to yield a memory score from 0 to 20
- Numeracy
 - Based on a set of questions on percentage calculation summarized in a score that ranges from 1 (good) to 5 (bad).
 - In waves 4 and 5, baseline respondents who already participated in one of panel waves are given a new test based on subtraction.
 - Correct answers lead to more difficult questions, while wrong answers lead to easier questions

Data - Measures of health

- Depression score (EURO-D)
 - The score is a sum of 12 binary indicators of whether the respondent experienced the following sentiments in the previous month: depressed mood, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment, tearfulness
 - The score ranges from 0 to 12

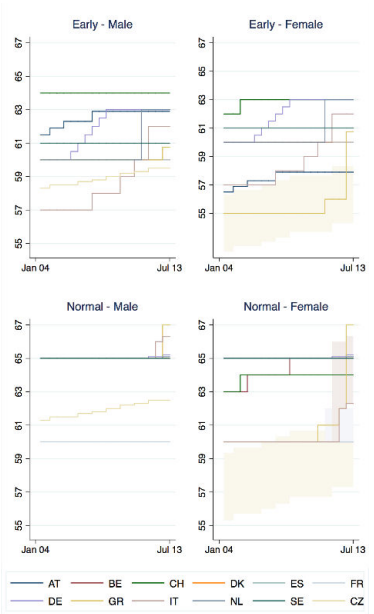
Data - Measure of work intensity

- Dummies for working part-time and full-time
 - Full-time work dummy indicates working 35 or more hours a week for 8 months or more in a year
 - Part-time work dummy indicates working less than 35 hours a week for 8 months or more a year, or working 35 or more hours a week but less than 8 months a year
 - The base group is retirement which is working 0 hours
 - The hours and months from both the main and a possible second job are considered

Data - Instruments for work status

- We use two sets of instruments for part-time and full-time work status
 - The first set includes two dummies which indicate whether the individual is between the early and normal retirement age, or at or above the normal retirement age
 - The second set includes two other dummies which indicate whether the partner is between the early and normal retirement age, or at or above the normal retirement age

Data - Variation in retirement eligibility ages



Data - Instruments for work status

Employment rates at the retirement eligibility ages (%)

Eligibility age	Full-time worker	Part-time worker	Full-time retiree
Under early ret. age	67.83	23.74	8.43
Between early and normal ret. age	33.70	11.85	54.45
Over the normal ret. age	6.51	3.08	90.41
Under early ret. age (P)	56.33	18.52	25.15
Between early and normal ret. age (P)	27.88	12.26	59.86
Over the normal ret. age (P)	10.91	5.37	83.72

Notes: 1. P: Partner. 2. Disabled, unemployed, and 'not in the labor force' are excluded from the analysis.

Data - Descriptive statistics

Descriptive statistics: demographics and work status

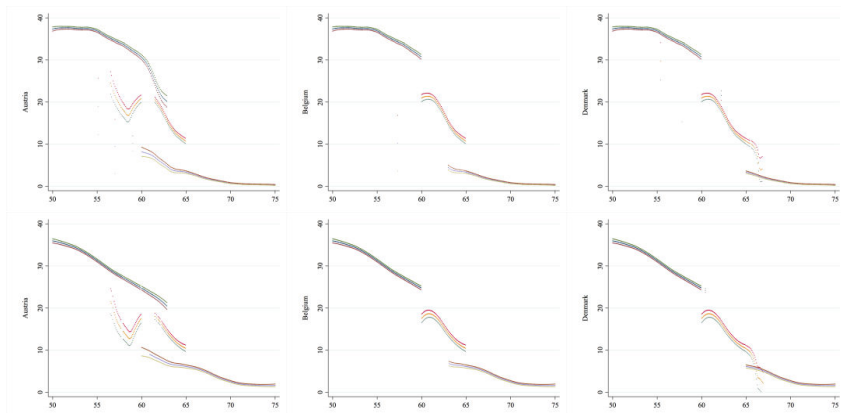
	Percent		
	All waves	2004 wave	2013 wave
Age (50-75) (avg.)	62.47	61.82	62.75
Under early ret. age	42.09	43.03	42.57
Between early and normal ret. age	14.27	16.31	13.78
Over normal retirement age	43.64	40.66	43.65
High education	28.89	24.56	31.73
Partner (married or unmarried)	80.09	79.01	79.96
Female	44.32	40.78	47.30
Full-time worker	33.64	34.61	33.76
Part-time worker	13.23	12.46	13.96
Full-time retiree	53.13	52.93	52.28
N obs.	70949	13390	25478
N ind.	41514	13390	25478

Data - Descriptive statistics

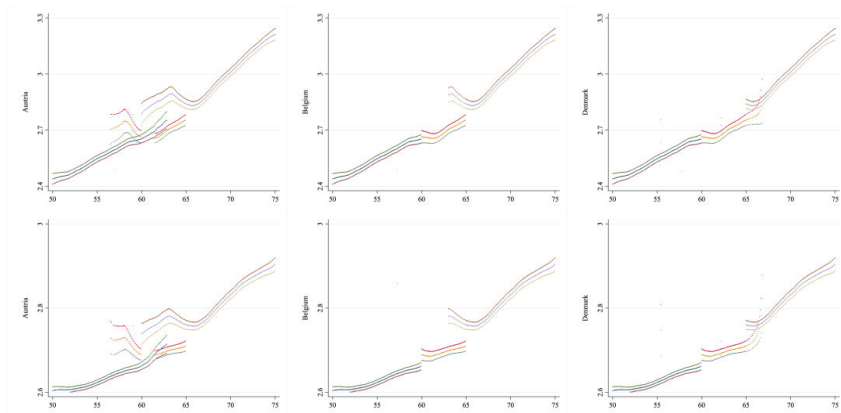
Descriptive statistics: health status

	Percent		
	All waves	2004 wave	2013 wave
Self-perceived fair or poor health	22.54	20.42	22.09
N of ADL limitations (0-5) (avg.)	0.08	0.07	0.08
N of IADL limitations (0-5) (avg.)	1.80	1.58	2.02
N of mobility limitations (0-5) (avg.)	0.25	0.26	0.23
N of diff. in muscle use (0-4) (avg.)	0.43	0.45	0.42
N of chronic diseases (0-9) (avg.)	0.90	0.89	0.91
Hospital stay in the previous two years	11.75	10.55	12.09
Overweight	42.94	44.87	41.55
Obese	16.63	15.84	16.82
Word recall test score (0-20) (avg.)	9.80	8.82	10.28
Numeracy (0-5) (avg.)	4.06	3.57	4.36
Depression scale EURO-D (0-12) (avg.)	1.89	1.86	1.90
Fluency (0-100) (avg.)	21.45	20.30	22.40
N obs.	70949	13390	25478
N ind.	41514	13390	25478

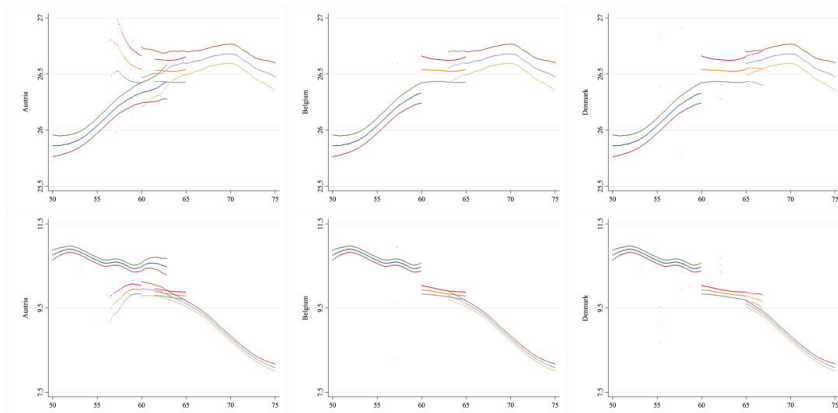
Data - Hours worked per week by age of respondent and partner allowing for jumps at the eligibility ages: kernel smoothed local polynomials, 95% CI around them



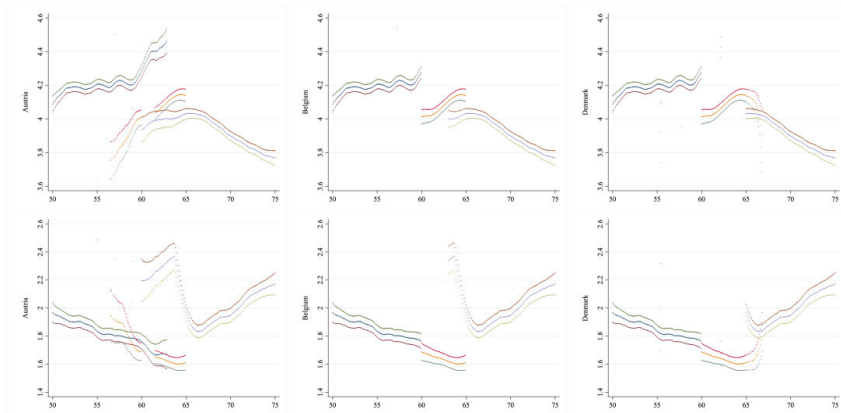
Data - Self-perceived health and health index by age of respondent allowing for jumps at the eligibility ages: kernel smoothed local polynomials, 95% CI around them



Data - Body mass index and word recall score by age of respondent allowing for jumps at the eligibility ages: kernel smoothed local polynomials, 95% CI around them



Data - Numeracy score and depression score by age of respondent allowing for jumps at the eligibility ages: kernel smoothed local polynomials, 95% CI around them



Results - Health outcomes

IV-FE model explaining health outcomes

	Self-perceived health		Health index		Body mass index	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	-0.714	0.379	-0.893	0.017‡	5.295	0.066‡
Full-time	0.572	0.041	0.371	0.004	-1.779	0.044
Age	0.047	0.000	0.015	0.000	0.030	0.006
Exo. test		0.000		0.000		0.001
Ove. test		0.143		0.516		0.439
F-test PT and FT		0.000		0.009		0.130
N obs.	43248		41871		25747	
N ind.	17518		16961		10697	
Model: Linear age, 4 IV, FE						

Notes: 1. Self-perceived health: 1 (very good), ..., 5 (very bad). Health index takes similar values. BMI takes values from 10.9 to 82.7. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Results - Health outcomes

IV-FE model explaining health outcomes

	Word recall score		Numeracy		Depression score	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	10.122	0.024 ‡	-2.373	0.063 ‡	-5.612	0.024 ‡
Full-time	-3.942	0.012	0.832	0.060	2.272	0.008
Age	0.015	0.417	0.093	0.000	0.042	0.000
Exo. test		0.000		0.040		0.000
Ove. test		0.861		0.241		0.393
F-test PT and FT		0.037		0.167		0.019
N obs.	42531		43069		42389	
N ind.	17215		17441		17170	
Model: Linear age, 4 IV, FE						

Notes: 1. Word recall score takes values from 0 to 20. Higher values indicate better memory. Numeracy takes values from 1 (bad) to 5 (good). Depression score takes values from 0 to 12. Higher values indicate more severe depression. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

First-stage FE model explaining part-time work status

	Part-time					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Bet. early and nor.	-0.041	0.000	-0.042	0.000	-0.030	0.000
At or over the nor.	-0.104	0.000	-0.104	0.000	-0.080	0.000
Bet. early and nor. (P)	0.005	0.420	-0.005	0.442	0.007	0.291
At or over the nor. (P)	-0.000	0.972	-0.000	0.978	0.002	0.756
Age	-0.004	0.000	-0.002	0.727	0.450	0.000
Age ²			0.000	0.671	-0.007	0.000
Age ³					0.000	0.000
Constant	0.440	0.000	0.372	0.044	-8.864	0.000
F-test 2 age terms				0.000		
F-test 3 age terms						0.000
F-test 4 instruments		0.000		0.000		0.000
AP test of weak ide.		0.000		0.000		0.000
Model: Linear probability model with FE						

Notes: 1. P: married or unmarried partner. 2. Standard errors are robust to heteroskedasticity and clustering on panel respondents.

Robustness checks - Age splines

First-stage FE model explaining full-time work status

	Full-time					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Bet. early and nor.	-0.148	0.000	-0.137	0.000	-0.088	0.000
At or over the nor.	-0.303	0.000	-0.301	0.000	-0.197	0.000
Bet. early and nor. (P)	-0.026	0.001	-0.021	0.011	-0.012	0.131
At or over the nor. (P)	-0.009	0.350	-0.010	0.275	-0.000	0.963
Age	-0.020	0.000	-0.072	0.000	1.886	0.000
Age ²			0.000	0.000	-0.031	0.000
Age ³					0.000	0.000
Constant	1.769	0.000	3.391	0.000	-36.594	0.000
F-test 2 age terms				0.000		
F-test 3 age terms						0.000
F-test 4 instruments		0.000		0.000		0.000
AP test of weak ide.		0.000		0.000		0.000
Model: Linear probability model with FE						

Notes: 1. P: married or unmarried partner. 2. Standard errors are robust to heteroskedasticity and clustering on panel respondents.

Robustness checks - Age splines

IV-FE model explaining self-perceived health

	Self-perceived health					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	-0.714	0.379	-0.231	0.825	-0.336	0.749
Full-time	0.572	0.041	0.405	0.265	0.499	0.253
Age	0.047	0.000	-0.001	0.955	-0.290	0.566
Age ²			0.000	0.124	0.005	0.544
Age ³					-0.000	0.579
Exo. test		0.000		0.002		0.015
Ove. test		0.143		0.076		0.090
F-test PT and FT		0.000		0.000		0.016
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. Self-perceived health: 1 (very good), ..., 5 (very bad). 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

IV-FE model explaining health index

	Health index					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	-0.893	0.017‡	-0.235	0.421	-0.211	0.465
Full-time	0.371	0.004	0.142	0.160	0.163	0.177
Age	0.015	0.000	-0.036	0.000	-0.199	0.178
Age ²			0.000	0.000	0.003	0.216
Age ³					-0.000	0.291
Exo. test		0.000		0.006		0.054
Ove. test		0.516		0.493		0.440
F-test PT and FT		0.009		0.012		0.095
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. Health index: 2 to 5. Higher values indicate worse health. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

IV-FE model explaining body mass index

	Body mass index					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	5.295	0.066‡	3.058	0.198	3.080	0.209
Full-time	-1.779	0.044	-1.150	0.115	-1.173	0.219
Age	0.030	0.006	0.296	0.000	0.375	0.803
Age ²			-0.002	0.000	-0.003	0.889
Age ³					0.000	0.958
Exo. test		0.001		0.065		0.250
Ove. test		0.439		0.883		0.884
F-test PT and FT		0.130		0.233		0.442
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. BMI takes values from 10.9 to 82.7. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

IV-FE model explaining word recall score

	Word recall score					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	10.122	0.024‡	3.991	0.338	3.609	0.384
Full-time	-3.942	0.012	-1.801	0.221	-2.355	0.177
Age	0.015	0.417	0.533	0.000	4.356	0.030
Age ²			-0.004	0.000	-0.066	0.047
Age ³					0.003	0.067
Exo. test		0.000		0.083		0.059
Ove. test		0.861		0.550		0.471
F-test PT and FT		0.037		0.233		0.151
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. Word recall score takes values from 0 to 20. Higher values indicate better memory. Higher values indicate worse health. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

IV-FE model explaining numeracy score

	Numeracy score					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	-2.373	0.063‡	-4.017	0.084‡	-3.852	0.097‡
Full-time	0.832	0.060	1.400	0.085	1.735	0.075
Age	0.093	0.000	0.247	0.001	-1.924	0.085
Age ²			-0.001	0.023	0.034	0.066
Age ³					-0.000	0.061
Exo. test		0.040		0.008		0.000
Ove. test		0.241		0.278		0.259
F-test PT and FT		0.167		0.221		0.202
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. Numeracy takes values from 1 (bad) to 5 (good). Higher values indicate worse health. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Age splines

IV-FE model explaining depression score

	Depression score					
	Linear age		Quadratic age		Cubic age	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	-5.612	0.024‡	-3.444	0.193	-3.141	0.218
Full-time	2.272	0.008	1.521	0.099	1.649	0.119
Age	0.042	0.000	-0.110	0.191	-1.450	0.247
Age ²			0.001	0.057	0.022	0.271
Age ³					-0.000	0.305
Exo. test		0.000		0.108		0.168
Ove. test		0.393		0.324		0.269
F-test PT and FT		0.019		0.087		0.219
F-test age terms				0.000		0.000
Model: 4 IV, FE						

Notes: 1. Depression score takes values from 0 to 12. Higher values indicate more severe depression. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Instrument set

IV-FE model estimated using a restricted instrument set

	Self-perceived health		Health index		Body mass index	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	0.037	0.971	-1.486	0.063‡	8.584	0.113
Full-time	0.286	0.424	0.580	0.043	-3.013	0.079
Age	0.043	0.000	0.015	0.000	0.021	0.124
Exo. test		0.000		0.000		0.000
Ove. test		—		—		—
Model: Linear age, 2 IV, FE						
Part-time	-0.714	0.379	-0.893	0.017‡	5.295	0.066‡
Full-time	0.572	0.041	0.371	0.004	-1.779	0.044
Age	0.047	0.000	0.015	0.000	0.030	0.006
Exo. test		0.000		0.000		0.001
Ove. test		0.143		0.516		0.439
Model: Linear age, 4 IV, FE						

Notes: 1. Self-perceived health: 1 (very good), ..., 5 (very bad). Health index takes similar values. BMI takes values from 10.9 to 82.7. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Instrument set

IV-FE model estimated using a restricted instrument set

	Word recall score		Numeracy		Depression score	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	10.900	0.105‡	-3.941	0.099‡	-10.304	0.082
Full-time	-4.291	0.076	1.462	0.085	3.878	0.066
Age	0.028	0.146	0.097	0.000	0.043	0.009
Exo. test		0.001		0.003		0.000
Ove. test		—		—		—
Model: Linear age, 2 IV, FE						
Part-time	10.122	0.024‡	-2.373	0.063‡	-5.612	0.024‡
Full-time	-3.942	0.012	0.832	0.060	2.272	0.008
Age	0.015	0.417	0.093	0.000	0.042	0.000
Exo. test		0.000		0.040		0.000
Ove. test		0.861		0.241		0.393
Model: Linear age, 4 IV, FE						

Notes: 1. Health index: 2 to 5. Higher values indicate worse health. 2. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Robustness checks - Econometric model

Econometric models explaining self-perceived health and depression score

	Self-perceived health		Depression score	
	Coeff	p-val	Coeff	p-val
Part-time	-0.276	0.000‡	0.058	0.030‡
Full-time	-0.367	0.000	-0.357	0.000
Model: Linear age, Pooled OLS				
Part-time	0.049	0.014	0.144	0.001
Full-time	0.026	0.121	0.165	0.000
Model: Linear age, FE				
Part-time	1.105	0.000‡	3.028	0.000‡
Full-time	-0.584	0.000	-1.395	0.000
Exo. test		0.000		0.000
Ove. test		0.000		0.000
Model: Linear age, Pooled, 4 IV				
Part-time	-0.714	0.379	5.612	0.024‡
Full-time	0.572	0.041	2.272	0.008
Exo. test		0.000		0.000
Ove. test		0.143		0.393
Model: Linear age, 4 IV, FE				

Notes: 1. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents.
 3. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Comparison with US

Health outcomes across US and Europe

	Self-perceived health		Health index		Body mass index	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	0.984	0.004‡	0.075	0.603	-3.197	0.004‡
Full-time	0.271	0.004	0.058	0.151	-0.935	0.002
Age	-0.004	0.785	0.004	0.536	0.582	0.000
Age ²	0.000	0.012	0.000	0.010	-0.004	0.000
Exo. test		0.000		0.000		0.000
Ove. test		0.146		0.050		0.954
F-test PT and FT		0.009		0.186		0.005
N obs.	60952		44061		60218	
N ind.	12521		10181		12433	
Model: Quadratic age, 6 IV, FE. Sample: US						
Part-time	-0.714	0.379	-0.893	0.017‡	5.295	0.066‡
Full-time	0.572	0.041	0.371	0.004	-1.779	0.044
Age	0.047	0.000	0.015	0.000	0.030	0.006
Exo. test		0.000		0.000		0.001
Ove. test		0.143		0.516		0.439
F-test PT and FT		0.000		0.009		0.130
N obs.	43248		41871		25747	
N ind.	17518		16961		10697	
Model: Linear age, 4 IV, FE. Sample: Europe						

Notes: 1. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents.

2. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Comparison with US

Health outcomes across US and Europe

	Word recall score		SRM/Numeracy		Depression score	
	Coeff	p-val	Coeff	p-val	Coeff	p-val
Part-time	2.144	0.139	1.838	0.000 ‡	0.142	0.817
Full-time	0.899	0.024	0.451	0.000	0.227	0.189
Age	0.332	0.000	0.052	0.010	-0.096	0.003
Age ²	-0.003	0.000	0.000	0.668	0.000	0.002
Exo. test		0.158		0.000		0.026
Ove. test		0.097		0.172		0.832
F-test PT and FT		0.063		0.000		0.187
N obs.	46546		56050		51428	
N ind.	10837		11857		11377	
Model: Quadratic age, 6 IV, FE. Sample: US						
Part-time	10.122	0.024 ‡	-2.373	0.063 ‡	-5.612	0.024 ‡
Full-time	-3.942	0.012	0.832	0.060	2.272	0.008
Age	0.015	0.417	0.093	0.000	0.042	0.000
Exo. test		0.000		0.040		0.000
Ove. test		0.861		0.241		0.393
F-test PT and FT		0.037		0.167		0.019
N obs.	42531		43069		42389	
N ind.	17215		17441		17170	
Model: Linear age, 4 IV, FE. Sample: Europe						

Notes: 1. SRM: Self-rated memory. 1. Standard errors and specification tests are robust to heteroskedasticity and clustering on panel respondents. 2. ‡ indicates rejection of the equality of the coefficients of part-time and full-time.

Conclusion

- We have analyzed the effect of working part-time and full-time on physical and mental health of elderly
- Individual heterogeneity and selection appears to be important in the analysis of the effects of working on health at the intensive margin
- Working part-time and full-time have significant and opposite effects on both physical and mental health
- This suggests that health in old age respond to number of work hours in a non-linear fashion
- We will investigate if these effects have an age gradient
- We will investigate the mechanisms behind the causal effects we find
- Health responses to working part-time and full-time differ across the United States and Europe