Netspar Network for Studies on Pensions, Aging and Retirement

Norma Coe

Hans-Martin von Gaudecker

Maarten Lindeboom

Jürgen Maurer

The Effect of Retirement on Cognitive Functioning

Discussion Paper 10/2009 - 044

October 27, 2009

The Effect of Retirement on Cognitive Functioning*

Norma B. Coe
Boston College, Center for Retirement Research

Hans-Martin von Gaudecker†
VU University Amsterdam and Netspar

Maarten Lindeboom
VU University Amsterdam, IZA, and Netspar

Jürgen Maurer
The RAND Corporation

27 October 2009

Abstract

Cognitive impairment has emerged as a major driver of disability in old age, with profound effects on individual well-being and decision-making at older ages. Decelerating its decline among the elderly is one of the main challenges for ageing societies. In the light of policies aimed at postponing retirement ages, an important question is whether retirement has an influence on the descend rate. Among the life style and psychosocial risk factors, intellectual stimulation has often been mentioned as a key factor in maintaining high levels of cognitive functioning. We use data from the HRS to estimate a model for the change in cognitive functioning. As retirement and cognitive functioning may be endogenously related, we use unexpected early retirement window offers to instrument for retirement behavior. These offers are legally required to be unrelated to the baseline health of the individual, and are significant predictors of retirement. While the simple OLS estimates show a negative relation between retirement and the rate of decline in various measures of cognitive functioning, instrumental variables estimates suggest that this may not be a causal effect. In particular, we do not find a clear relationship for white-collar workers and a positive relation for blue-collar workers.

*We would like to thank seminar and conference participants at the University of Mannheim, Stockholm University, the RAND Corporation, and the 7^{th} iHEA World Congress in Beijing for very helpful comments.

†Corresponding author: VU University Amsterdam, Department of Economics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel: +31 20 598 3616

1 Introduction

Over the last century, longevity in the United States has increased from 47.3 years in 1900 to 77.8 years in 2005 (National Center for Health Statistics 2009). These gains in life-expectancy have been accompanied by large changes in morbidity patterns. The historically more important burden of infectious diseases has been largely replaced by degenerative diseases, which are largely concentrated among older persons. Particularly, cognitive impairment has emerged as a major driver of disability in old age, and its relative share of the total disease burden is likely to increase further as longevity continues to increase. Estimates suggest that the number of people with dementia is likely to roughly double every 20 years, resulting in a projected number of 81.1 million persons with dementia by 2040 (Ferri et al. 2005).

While some form of cognitive decline appears to be an inevitable affliction of old age, its progression can be relatively slow and might not result in any impairment of a person's wellbeing or ability to function on a daily basis. More serious forms of cognitive impairment may, however, have profound adverse effects on various aspects of life including leisure time activities , financial planning for retirement (Banks and Oldfield 2007) and functional status, such as medical treatment adherence and the planning of sequential activities (Fillenbaum et al. 1988). The list of potential risk factors for increased rates of memory loss is long and diverse, ranging from genetic factors over medical comorbidities to life-style and psychosocial factors (Institute for the Study of Aging 2001). Identifying such risk factors and developing strategies to maintain high levels of cognition have thus emerged as key public health priorities associated with population aging.

Among the life-style and psychosocial risk factors, continued intellectual stimulation is often seen as a key factor for retaining high levels of cognitive functioning at older ages (Small 2002). Specifically, cognitive training and intellectual stimulation may help maintain brain plasticity and thus prevent accelerated memory loss at older ages. Suggested strategies aimed at delivering continued intellectual stimulation include formal training as well as informal training activities, such as crossword puzzles or cognitively challenging social interactions.

Continued professional activities have also been suggested as potentially protective strategies against cognitive decline. For example, using data on older men from the Twins Registry of World War II veterans, Potter, Helms, and Plassman (2008) show that intellectually demanding work is positively associated with cognitive performance in later life. Similarly, Adam et al. (2007) use data from the Survey of Health, Ageing and Retirement in Europe (SHARE), the English Longitudinal Survey on Ageing (ELSA) and the U.S. Health and Retirement Study (HRS) to show a strong negative association of retirement and cognition for both older European and Americans, and that the strength of this association is monotonically increasing with length of retirement.

While the above associations between continued employment and cognitive functioning are clearly suggestive, they have to be taken with a grain of salt, as potentially important endogeneity issues caution against a causal interpretation. The positive association between continued employment and cognitive function may, of course, at least in part be due to continued intellectual stimulation on the job. It may, however, be further exacerbated by potential reverse causation, as maintaining a certain level of cognition is likely to be a necessary condition for gainful employment, especially for intellectually demanding jobs. Identifying a causal effect of continued employment on cognitive functioning thus requires exogenous variation in job retention to rule out potentially confounding effects from cognition on employment.

An additional complication stems from the fact that we should like to exploit exogenous variation in employment that is not a deterministic function of age. Specifically, the impact of intellectual inactivity is likely to be cumulative over time. We thus require continuous exogenous variation in retirement that is not functionally dependent on age in order to estimate the cumulative effects of retirement on cognitive functioning, controlling for potentially confounding age patterns of cognitive decline. This requirement prevents us from using age-related discontinuities in retirement incentives as instruments, as is often done in related research that studies the instantaneous effects of retirement on different outcomes of interest (see for example Charles (2002) and Neuman (2008) for applications to mental and physical health, respectively).

This paper aims at identifying a causal effect of discontinued on-the-job stimulation on later-life cognition using early retirement windows as an instrument for labor market exits. Specifically, we develop a stochastic model for the change in cognitive functioning and derive the exclusion restrictions required for a valid instrument in the context of this model. We argue that the offering of an early retirement window will satisfy these conditions. We describe the results using various estimation strategies for several measures of cognitive functioning and discuss the findings at the end of the paper. We start out by describing the data we use.

2 Data

The data are from the Health and Retirement Study, conducted by the Survey Research Center at the University of Michigan. This is a longitudinal survey of people aged 50 and older and their spouses (regardless of age) starting in 1992 with follow-up interviews every 2 years. Currently, 8 waves of data are available (1992-2006). When weighted to account for initial over-sampling of some population groups and for subsequent attrition, the HRS provides a representative sample of the relevant birth cohorts. We use the RAND HRS data files (St. Clair et al. 2009) for all background variables and construct the variables on cognitive functioning from the raw HRS files based on Ofstedal, Fisher, and Herzog (2005). The latter are available from wave 3 onwards and we thus use data from all waves between 1996 and 2006.

2.1 Explanatory variables

We distinguish between white-collar and blue-collar workers in our analyses and report the descriptive statistics separately in Table 1. In order to complete the analysis, we make a variety of sample restrictions. First, we restrict the analyses to males because of selection issues associated with female labour force participation. Second, we limit the sample to those below 80 years of age and the main respondents belonging to the original HRS cohort (1931-1941) or younger. These age restrictions strike a balance between including ages with substantial declines in cognitive functioning on the one hand; and minimising reporting errors with respect to variables asked retrospectively in the first wave of interviews on the other hand. We experimented with various other cutoffs and did not detect differences. Selective mortality may also be an issue in the selection of our sample since we include individuals up to the age of 80. We return to this issue in Section 4.

The sample includes 23,615 person-year observations. Of these, 52% are for blue-collar workers as defined by their job with the longest tenure. The independent variables we include consist of wave and cohort dummies, age and age squared, and indicator variables for education (in five categories), race, and ethnicity. Some of our specifications contain interactions of these variables and we ran specification including various other variables, without finding different results.

We define retirement in two ways. First, we use a self-report on whether individuals consider themselves to be (partly) retired. The first panel in Figure 1 plots this variable separately for blue collar and white collar workers. The strongest rise can be seen for blue collar workers at age 62 when it is first possible to draw old age social security benefits. For white collar workers, the increase is more steady, but also strongest during individuals' early and mid sixties. By age 70, close to 90% of both types of workers are retired by this criterion. The second variable we use to define retirement is whether individuals are currently working for pay. Although not shown here, the age gradient is very similar to the one observed for the self-reports. However, it levels off at approximately 80%, which is likely to be mainly due to very small jobs. In the empirical specification, we use the time elapsed since the reported retirement date and the time elapsed since the date that the last job ended, respectively. Both of these variables are available in the data at the monthly level.

Table 1: Descriptive statistics by worker type

White Collar Workers

Variable	Obs	Mean	Std. Dev.	Min	Max
Immediate Recall	11206	5.913707	1.590847	0	10
Delayed Recall	11180	4.912433	1.836147	0	10
Total Recall	11180	10.83766	3.172224	0	20
Serial 7's	11212	4.308866	1.123353	0	5
Numeracy	3687	1.844318	.7968737	0	3
Self-Rated Memory	11264	3.211648	.9059194	1	5
Retired (self-reported)	11272	.578868	.4937626	0	1
Retired (not currently working for pay)	11272	.3975337	.4894098	0	1
Ever offered an Early Retirement Window	11272	.1936657	.3951875	0	1
Age	11272	64.125	6.233692	50	80
GED	11272	.0310504	.1734616	0	1
High-School Graduate	11272	.2011178	.4008537	0	1
Some College	11272	.240951	.4276796	0	1
College and Above	11272	.4667317	.4989141	0	1
Race: Black	11272	.0707062	.2563448	0	1
Race: Other	11272	.0282115	.1655839	0	1
Hispanic	11272	.0395671	.1949484	0	1

Blue Collar Workers

Variable	Obs	Mean	Std. Dev.	Min	Max
Immediate Recall	12227	5.103541	1.613905	0	10
Delayed Recall	12162	4.043743	1.879232	0	10
Total Recall	12162	9.171929	3.214702	0	20
Serial 7's	12033	3.494723	1.663115	0	5
Numeracy	3582	1.27694	.8558936	0	3
Self-Rated Memory	12334	2.862494	.9504376	1	5
Retired (self-reported)	12343	.6244835	.4842756	0	1
Retired (not currently working for pay)	12343	.5039294	.5000048	0	1
Ever offered an Early Retirement Window	12343	.1085636	.3111034	0	1
Age	12343	64.11221	6.251155	50	80
GED	12343	.0772908	.2670631	0	1
High-School Graduate	12343	.3588269	.4796757	0	1
Some College	12343	.1666532	.372681	0	1
College and Above	12343	.0626266	.2422999	0	1
Race: Black	12343	.1715952	.3770435	0	1
Race: Other	12343	.0443166	.2058059	0	1
Hispanic	12343	.1162602	.3205497	0	1

2.2 Cognitive Functioning

Our analysis exploits a series of measures for cognitive function as outcome variables. Specifically, we consider measures of self-rated memory, immediate, delayed and total word recall, working memory, and numeracy. These measures aim at covering different aspects of cognitive function within the framework of a large population-based general-purpose survey. A more detailed description of these survey instruments as well as rationale for their choice can be found in Ofstedal, Fisher, and Herzog (2005), on which this section draws.

Self-rated memory is assessed based on the survey instrument "How would you rate your memory at the present time? Would you say it is excellent, very good, good, fair, or poor?" This survey instrument mainly characterizes overall memory perceptions of respondents rather than "objective" memory status. Yet, over and above being a useful measure of memory perceptions, this measure has also been shown to be correlated with more objective measures of memory performance (Maurer 2009). We translated verbal responses to the self-rated memory question to a quantitative scale ranging from 1 ("poor") to 5 ("excellent").

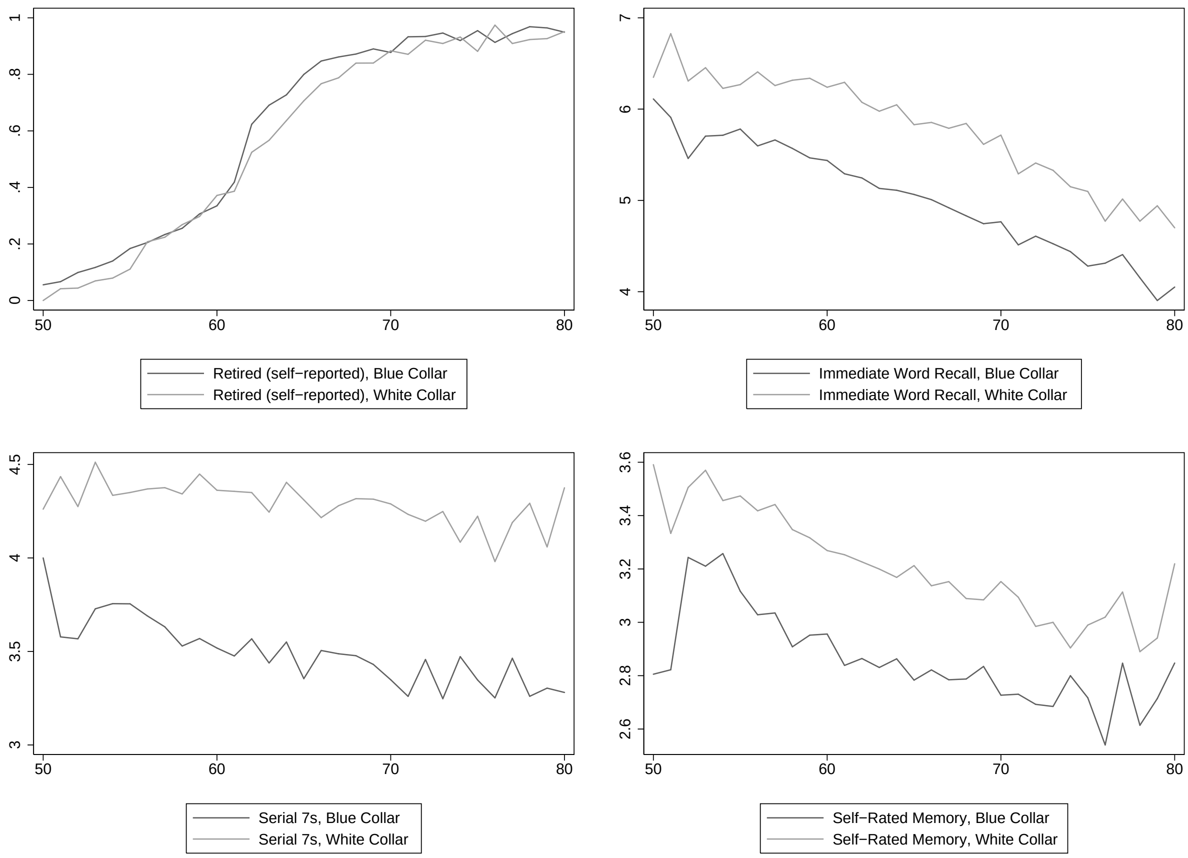
We contrast our findings for self-rated memory with evidence from more objective test-based performance measures of cognitive functioning. Immediate and delayed word recall aim at assessing memory performance based on two word recall tasks. Respondents were read a list of ten nouns and asked to recall as many words as possible from the list in any order. The score for immediate word recall counts the number of correct responses, leading to a test score between 0 and 10. Roughly five minutes later (after the administration of some additional survey instruments), the respondent is again asked to recall as many words from the previously read list of nouns. Again, corresponding test score is obtained as the sum of each correct answer (delayed work recall, range 0-10). We also combined the scores from both immediate and delayed word recall to obtain an overall summary measure for recall (total word recall), whose score ranges from 0 to 20.

Working memory, i.e. the ability to process and store information simultaneously, is assessed based on a serial 7's subtraction test. In the serial 7's test, respondents are asked to subtract 7 from 100 and continue subtracting 7 from each subsequent number for a total of five trials. This test thus requires respondents to perform a basic arithmetic operation (subtracting 7), while memorizing the result from the previous subtraction that is required as an input in this process. The serial 7's subtraction test score counts each correct subtraction, leading to scores between 0 and 5.

Finally, numeracy is measured as the number of correct responses to three numerical problems. The tasks involve 3 calculation tasks, i.e., calculation of percentages, division, and compound interest. Scoring correct solutions for each task yields a numeracy score ranging from 0 to 3.

Table 1 provides basic summary statistics for all these variables, separately for white and

Figure 1: Age trends in key variables



blue collar workers. There is substantial variation in all measures across the whole range of potential outcomes. Figure 1 shows that a sizable fraction of this variation is indeed related to age. It contains the age-related evolution of immediate recall, serial 7's, and self-rated memory, also broken down by worker type. As would be expected, white collar workers have a substantially higher level of cognitive functioning at every age. With the exception of sample-size related fluctuations at the boundaries of the selected age groups, the decline in all measures is largely parallel for both worker types. It is most pronounced for immediated recall and self-rated memory and much smaller for serial 7 substraction.

2.3 Early Retirement Windows

Early retirement windows are a limited-time offer, typically lasting six weeks to three months (Towers Perrin 1992). The HRS question starts with defining an early retirement window, stating:

> “Employers sometimes encourage older workers to leave a firm at a particular time by offering a special financial incentive, like a cash bonus or improved pension benefits. These are often called ‘early retirement windows’.”

The respondents are then asked:

> “Have you ever been offered such an early retirement window on any job?”

The survey then solicits information on how many offers the individual has received, when the offer(s) was received, which employer made it, what was offered in the plan, and whether it was accepted. If the offer was accepted, they ask if the offer was influential in their decision to leave the job. If the offer was rejected, they ask whether the offer would have induced leaving the job if the offer were doubled. In total 19% of white-collar workers and 10% of blue-collar workers were offered a retirement window (Table 1). Coe and Lindeboom (2008) analyse several of these variables in detail and conclude that the offered windows are a major factor in determining retirement decisions.

Coe and Lindeboom (2008) argue that conditional on observed characteristics, the offering of early retirement windows is exogenous to an individual’s health because firms cannot limit eligibility to single individuals. Instead, firms can only select broad groups of workers to be eligible for early retirement windows and their power to limit eligibility is bounded tightly by the courts. This reasoning also holds for cognitive functioning – while it is in the firms’ best interest to select individuals who experienced the greatest productivity decline, they do not have the power to do so. This is confirmed by statistical evidence. Exploring the panel nature of the data, Coe and Lindeboom (2008) do not find health status to predict offerings of retirement windows, conditional on observables. We performed similar analyses and did not find cognitive functioning to be a predictor of offerings of early retirement windows, either. We conclude from this that the exogeneity condition for the instrument – which we will make precise in the subsequent section – can be expected to hold in our sample. See Coe and Lindeboom (2008) for more specifics on the institutional features, details of the statistical analyses, and further references.

We operationalise the instrument by creating a set of dummy variables containing indicators of the age at the first offering of an early retirement window (until age 49, and annual indicators thereafter until age 65). We experimented with different specifications and did not find qualitative differences. By the rule of thumb due to Staiger and Stock (1997), there is no sign of weak instruments in most of our regressions. However, the F-tests are smaller than 10 for several of the specifications using the “not currently working for pay” variable to define retirement. It is for this reason that we prefer the self-reported retirement indicator and perform robustness checks with alternatives to two stage least squares.

3 Model and Empirical Specification

When it comes to cognition, it is hard to conceive of a Grossman (1972) - type model where individuals choose their optimal level of cognitive functioning. One reason is that the literature has focused on correlates for the onset of dementia and not on detecting causal effects (see the discussion in the Introduction). So there is little guidance to specify an investment model because there hardly is a priori information on inputs to the production function. We take a different route instead and focus on a statistical model for cognitive functioning, which is general enough to capture the mechanisms we are interested in.

3.1 Statistical Model for the Change in Cognitive Functioning

The literature in empirical labour economics has gained substantially in the past three decades from using time-series processes to model individual earnings dynamics (see MaCurdy (1982) for an early example and Meghir and Pistaferri (2004) for a recent extension and overview). Adda, Banks, and von Gaudecker (forthcoming) apply this framework to a variety of health indicators, focussing on the decomposition of transitory and permanent shocks. Doing so is fruitful for cognitive functioning as well because the indicators we use are not immune to many short-term factors associated with concentration on a particular day etc. – one may think about part of the transitory shocks in terms of measurement error.

We denote the age t change in cognitive functioning by ΔCF_t and decompose it into a permanent (ρ_t) and a transitory (ε_t) component:

$$\Delta CF_t = \rho_t + \Delta\varepsilon_t \tag{1}$$

We focus on the permanent component and model its mean as a linear function of observed covariates:

$$\rho_t = \alpha_t + RET_t \cdot \gamma + X_t\beta + \psi_t \tag{2}$$

Our parameter of interest is γ, i.e. the effect of being retired at age t on permanent innovations to cognitive functioning. This specification restricts γ to be the same across ages. While this assumption is certainly debatable and could be relaxed in theory, doing so is infeasible empirically for us and will have to await better datasets to become available. One may also view γ to be the average of a set $\gamma_t,\ t \in \{0, 1, \ldots, T\}$ with possibly different elements. Since we focus who have worked at some point, the model is completed by the following initial conditions:

$$CT_0 = \alpha_0 + X\beta + \psi_0 \tag{3}$$

3.2 Empirical Specification

Previous research using time-series approaches to changes in wages or health has sought to estimate the entire dynamic structure of (1). Since we do not have a sufficiently long panel available for such a formidable task, our aims are more modest. In order to develop an equation that we can estimate with our data, we first derive the cross-sectional implications of the model (1)-(3):

$$CF_t = \sum_{s=0}^{t} \alpha_s + \gamma \sum_{s=0}^{t} RET_s + \sum_{s=0}^{t} X_s \beta + \sum_{s=0}^{t} \psi_s + \varepsilon_t - \varepsilon_0 \tag{4}$$

From (4), it is evident that it is *not* enough to have contemporaneous covariates in the model, but that the entire history is important. It is for this reason that we focus on covariates that are stable over time. In order to identify γ, we first assume ψ_t and ε_t to be independent across time and individuals, as is common in the literature. Furthermore, we need the following conditions to hold for instrumenting $\sum_{s=0}^{t} RET_s$, the time spent in retirement until age t, with a variable Z_t:

1. $\theta \neq 0$ in $\sum_{s=0}^{t} RET_s = a_t + Z_t\theta + \sum_{s=0}^{t} X_s \cdot b + u_t$
2. $cov\left[Z_t, \sum_{s=0}^{t} \psi_s + \varepsilon_t - \varepsilon_0\right] = 0$

The first condition implies that conditional on age there has to be non-trivial variation in the instrument. This rules out using ages 62 and 65 – where retirement incentives induced by the social security system are strongest – as instruments, an approach taken by previous authors in order to estimated the impact retirement has on health (Charles 2002, Neuman 2008). We hence use the offering of early retirement windows, similar to Coe and Lindeboom (2008). As discussed in Section 2.3, the instruments are strong enough in the sense of Staiger and Stock (1997) in our preferred specification, so the first condition is satisfied in our data. The second condition is the instrument exogeneity condition and says that time elapsed since the first offering of an early retirement window may not be correlated with the (sum of the) unexplained permanent innovations to cognitive functioning, or the difference between the initial and contemporaneous transitory innovations. We provided both institutional and statistical reasons in Section 2.3 for why we expect this condition to hold.

In order to estimate (4), we work with monthly data and use pooled cross-sections of all waves between 1996 and 2006. Heteroskedasticity-robust standard errors are adjusted for clustering at the individual level. The gains from using this approach instead of a single cross-section come from reducing the impact of ε_t and from a modest increase in the clustering-adjusted sample size. Z_t is a vector of indicator variables containing the individual's age when the first retirement window was offered. In our preferred specification, we pool blue- and white-collar workers but interact instruments, time elapsed since the retirement date, and the quadratic age trend with the indicator for worker type. X_t furthermore includes education, race, ethnicity, and wave dummies. Results from separate regressions by worker type and specifications including various interaction terms and/or cohort dummies were not significantly different from these estimates and are available upon request.

Another robustness check we include is to specify the model in differences between waves 2006 and 1996, which is straightforward from (4). This reduces both our sample size and the number of variables we can use by substantial amounts. Because our instruments show signs of being weak if we use work status ("currently working for pay") as the retirement variable we implement the estimators described in Hausman et al. (2008) as alternatives.[1] These estimators are inconsistent in clustered samples, so we only use the 2006 wave of the data for estimation in this case.

4 Description of Results

Our main results are contained in Table 2. First consider the rows 1 and 6 containing the OLS coefficients. Similar to Adam et al. (2007), we find the expected negative associations between length of retirement and the measures of cognitive functioning for both white- and blue-collar workers. Only the serial 7's for white collar workers show the opposite sign, but the coefficient is far from being significant. In contrast, most of the negative coefficients are significant at the 5%-level. Taking the first element of the table as an example, the interpretation of the coefficients is such that an additional month spent in retirement is associated with an average reduction of .00126 correctly recalled words (out of 10). These effects are small: An additional ten years in retirement are associated with 0.15 less correctly remembered words, holding everything else (age, in particular) constant.

The direction of these associations continues to hold for white collar workers when we move to the instrumental variables (IV) estimation in the third row of Table 2. As often the case with instrumental variables, most coefficients increase in absolute value and so do their standard errors (Angrist and Krueger 1991, Grimard and Parent 2007). Only the numeracy coefficient switches sign and its standard error is so large that we cannot distinguish it from the OLS coefficient. In fact, it is only for self-rated memory that we can reject the null hypothesis that the instrumental variables coefficient is the same as the least squares coefficient (test statistics are in row 5). In a local average treatment effects interpretation (Imbens and Angrist 1994), the coefficient implies that a 10-year additional retirement spell induced by the offering of an early retirement window leads to a .58 point reduction in self-rated memory, as measuered on a 5-point scale.

The picture is quite different for blue-collar workers. All IV estimates now indicate a

[1] Note that the simple limited-information maximum likelihood estimator is inconsistent under weak instruments and heteroskedasticity, where the latter is implied by (4).

Table 2: Results from the preferred specification – pooled cross sections and self-reported retirement

		Immediate Recall	**Delayed Recall**	**Total Recall**	**Serial 7's**	**Numeracy**	**Self-Rated Memory**
White Collar	OLS	-0.00126	-0.00145	-0.00270	0.00034	-0.00043	-0.00084
	S.E.	0.00045	0.00055	0.00095	0.00034	0.00028	0.00028
	2SLS	-0.00178	-0.00019	-0.00182	0.00152	0.00063	-0.00484
	S.E.	0.00309	0.00375	0.00652	0.00238	0.00165	0.00222
	p(OLS=2SLS)	0.844	0.691	0.874	0.563	0.434	0.039
Blue Collar	OLS	-0.00084	-0.00056	-0.00132	-0.00115	-0.00031	-0.00066
	S.E.	0.00035	0.00042	0.00072	0.00039	0.00025	0.00025
	2SLS	0.01250	0.00763	0.01943	0.00616	0.00477	0.00482
	S.E.	0.00569	0.00593	0.01091	0.00587	0.00343	0.00301
	p(OLS=2SLS)	0.012	0.137	0.042	0.181	0.109	0.047
N	total	22228	22155	22155	22075	7599	22372
	clustered	6083	6075	6075	6066	4736	6100

Note: Regressions with pooled blue- and white-collar workers. The instruments, retirement status, and the quadratic age trend are interacted with the indicator for worker type. Further controls includes education, race, ethnicity, and wave dummies. Data are pooled cross-sections of all waves between 1996 and 2006. Heteroskedasticity-robust standard errors are adjusted for clustering. The rows labeled p(OLS=2SLS) contain p-values for a Hausman test for the difference between OLS and 2SLS estimates (under the null hypothesis of OLS being unbiased and efficient, $(\gamma_{j,\text{OLS}} - \gamma_{j,\text{2SLS}})^2/(\sigma_{j,\text{2SLS}} - \sigma_{j,\text{OLS}})^2 \sim \chi^2_{(1)}$)

positive effect of retirement on cognitive functioning. For immediate and total recall the coefficients are significant at the 5%-level, for self-rated memory it is almost significant at the 10%-level. The magnitudes of the estimated effects are substantial – a 10-year additional retirement spell induced by the offering of an early retirement window leads to a 1.5 word decrease in the number of correctly remembered items. This would more than close the gap between blue- and white collar workers. It has to be noted though that the standard errors are such that they do not exlude smaller effects. However, we can reject the null hypothesis of OLS and 2SLS being the same in three of the six cases, in the remaining columns the p-value is between 10% and 18%.

The general picture of an unclear or negative effect of retirement on cognitive functioning for white-collar workers and of a positive such effect for blue-collar workers also emerges in our various other specifications. In particular, the pattern is robust to a large number of interaction terms among the explanatory variables and to separate estimations by worker type. In the appendix, we include the tables corresponding to Table 2 for the model in 2006-1996 differences (Table 3) and the one with retirement time defined by the end date of the last job (Table 4). Estimates from the first difference model will take selective mortality into account as long as mortality is driven by time constant individual factors. The reduction in sample size by two thirds for the difference specification in Table 3 mainly compromises the significance of results, although the coefficient on the serial 7's now exhibits a negative sign for both worker types (albeit with very large standard errors). The significant estimates in Table 4 have the same pattern as before. However, this does not hold up for all other coefficients anymore. The reason behind this is likely to be the fact mentioned earlier, namely that our instrument is weak for the alternative definition of the retirement variable. We ran the regressions for the 2006 wave only using the heteroskedasticity-robust version of the limited information maximum likelihood and Fuller estimators developed by Hausman et al. (2008). Coefficients and standard errors once more increased substantially in magnitude when using a single wave, both for 2SLS and the alternatives. Most likely the remaining sample is too small to render the asymptotic results to be a good approximation of the estimators' behaviour.

5 Discussion

Our estimation results highlight a negative association between length of retirement and most measures of cognitive functioning. These results obtain for both white and blue collar workers. Albeit quantitatively small, the negative associations are generally statistically significant, and thus suggest a potential negative relationship between time spent in retirement and cognition.

Yet, robust empirical evidence for a negative gradient of cognitive functioning in retirement needs to account for potential reverse causation, which may stem from differential selection into retirement based on cognitive ability. To this end, we advanced an instrumental variables strategy that isolates exogenous variation in the length of retirement based on the offering of early retirement windows by employers.

Our instrumental variables estimates strongly caution against a causal interpretation of the negative association between retirement duration and cognitive function. Specifically, while we continue to find a strong and statistically significant negative gradient of self-rated memory in retirement for white collar workers (very much in line with anecdotal evidence concerning the potentially adverse effects of retirement on cognition), we find no systematic evidence for such a relationship based on our IV models for objective cognition measures, which is – admittedly – partially attributable to a lack of precision in the estimations. In this respect our findings are very similar to those obtained for other health variables. Charles (2002) Neuman (2008), Bound and Waidmann (2007), and Coe and Lindeboom (2008) also find that the strong and significant negative associations between various health measures and retirement do not hold when retirement is instrumented.

Importantly, our IV results for blue collar workers differ substantially from those for white collar workers discussed above. Specifically, our initial evidence for a negative relationship between cognition and retirement duration turns into a positive one for our IV estimates for all measures of cognitive function we consider. Moreover, while we are still unable to obtain very precise estimates using instrumental variables, we do obtain a statistically significant positive relationship between several cognition variables and retirement for blue collar workers. When compared to our OLS results, we obtain a statistically different coefficient for half of our measures of cognitive functioning. Estimated differences between the OLS and IV coefficients for our remaining measures are also suggestive for a sign reversal.

Overall, our results show that the negative association between cognitive function and length of retirement needs to be interpreted with caution. While our IV estimates for white collar workers lead to mixed evidence at best, corresponding regressions for blue collar workers suggest that OLS regressions overstate a potential negative relationship between cognition and retirement. In fact, our IV regressions for blue collar workers rather indicate that retirement has a positive effect for these workers, potentially reflecting a lack of mental challenges on the job for these workers (and relatively higher cognitive stimulation during retirement). Clearly, our approach of isolating exogenous variation in retirement length while flexibly controlling for simultaneous age effects is very ambitious, and we do pay a price for such flexible modeling in terms of estimation precision. Nevertheless, we deem this flexibility to be very important and the HRS is the best dataset currently available for the questions addressed in this paper. The results clearly indicate that simple OLS regressions will – in general – not be informative on the effects of retirement duration on cognitive functioning. We thus consider our study an important first step toward modeling these effects within an instrumental variables framework, even if further research will be needed to obtain more precise estimates of these effects.

References

Adam, S., E. Bonsang, S. Germain, and S. Perelman (2007): "Retirement and Cognitive Reserve: A Stochastic Frontier Approach Applied to Survey Data," University of Liège, Center of Research in Public Economics and Population Economics Working Paper 2007/04.

Adda, J., J. Banks, and H.-M. von Gaudecker (forthcoming): "The Impact of Income Shocks on Health: Evidence from Cohort Data," *Journal of the European Economic Association.*

Angrist, J. D., and A. B. Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106(4), 979–1014.

Banks, J., and Z. Oldfield (2007): "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving," *Fiscal Studies*, 28(2), 143–170.

Bound, J., and T. Waidmann (2007): "Estimating the Health Effects of Retirement," University of Michigan, Retirement Research Center Working Paper 2007-168.

Charles, K. K. (2002): "Is Retirement Depressing? Labor Force Inactivity and Psychological Well-Being in Later Life," NBER Working Paper 9033.

Coe, N. B., and M. Lindeboom (2008): "Does Retirement Kill You? Evidence from Early Retirement Windows," IZA Discussion Paper No. 3817.

Ferri, C. P., M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang, A. Jorm, C. Mathers, P. R. Menezes, E. Rimmer, and M. Scazufca (2006): "Global Prevalence of Dementia: A Delphi Consensus Study," *The Lancet*, 366(9503), 2112–2117.

Fillenbaum, G. G., D. C. Hughes, A. Heyman, L. K. George, and D. G. Blazer (1988): "Relationship of Health and Demographic Characteristics to Mini-Mental State Examination Score among Community Residents.," *Psychological Medicine*, 18(3), 719–726.

Grimard, F., and D. Parent (2007): "Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking?," *Journal of Health Economics*, 26(5), 896–926.

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

Hausman, J., W. K. Newey, T. Woutersen, J. Chao, and N. Swanson (2008): "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments," CEMMAP Working Paper 22-07, updated version.

Imbens, G. W., and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–475.

Institute for the Study of Aging (2001): "Achieving and Maintaining Cognitive Vitality with Aging," Workshop Report, available at http://www.aging-institute.org.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.

Maurer, J. (2009): "Forgotten Forgetfulness? Memory Perceptions vs. Memory Performance of Older Americans," Mimeo, The RAND Corporation.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

National Center for Health Statistics (2009): *Health, United States, 2008.* Hyattsville, MD.

Neuman, K. (2008): "Quit Your Job and Get Healthier? The Effect of Retirement on Health," *Journal of Labor Research*, 29(2), 177–201.

Ofstedal, M. B., G. G. Fisher, and A. R. Herzog (2005): "Documentation of Cognitive Functioning Measures in the Health and Retirement Study," Survey Research Center, University of Michigan, HRS Documentation Report DR-006.

Potter, G. G., M. J. Helms, and B. L. Plassman (2008): "Associations of Job Demands and Intelligence with Cognitive Performance among Men in Late Life.," *Neurology*, 70(19 Pt 2), 1803–1808.

Small, G. W. (2002): "What We Need to Know about Age Related Memory Loss," *British Medical Journal*, 324, 1502–1505.

St. Clair, P., D. Blake, D. Bugliari, S. Chien, O. Hayden, M. Hurd, S. Ilchuk, F.-Y. Kung, A. Miu, C. Panis, P. Pantoja, A. Rastegar, S. Rohwedder, E. Roth, J. Carroll, and J. Zissimopoulos (2009): "RAND HRS Data Documentation, Version I," RAND Labor and Population Program, RAND Center for the Study of Aging.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(3), 557–586.

Towers Perrin (1992): "More Employers Opening Retirement Windows," *Monitor Newsletter*, 134, 1–3.

A Additional Tables

Table 3: Results based on 2006–1996 differences and self-reported retirement

		Immediate Recall	**Delayed Recall**	**Total Recall**	**Serial 7's**	**Self-Rated Memory**
White Collar	OLS	-0.00104	-0.00040	-0.00152	-0.00069	-0.00123
	S.E.	0.00097	0.00108	0.00187	0.00078	0.00051
	2SLS	-0.02763	-0.00027	-0.01827	-0.00852	0.00208
	S.E.	0.01876	0.00993	0.01574	0.01013	0.01051
	p(OLS=2SLS)	0.135	0.988	0.227	0.402	0.741
Blue Collar	OLS	-0.00079	0.00115	0.00071	-0.00165	0.00054
	S.E.	0.00095	0.00098	0.00167	0.00078	0.00054
	2SLS	0.03863	0.03030	0.07112	-0.00379	0.01002
	S.E.	0.02071	0.01685	0.03796	0.01270	0.00882
	p(OLS=2SLS)	0.046	0.066	0.052	0.858	0.253
N	clustered	1940	1932	1932	1927	1966

Note: Regressions with pooled blue- and white-collar workers. The instruments, retirement status, and the quadratic age trend are interacted with the indicator for worker type. Further controls includes education, race, ethnicity, and wave dummies. Data are pooled cross-sections of all waves between 1996 and 2006. Heteroskedasticity-robust standard errors are adjusted for clustering. The rows labeled p(OLS=2SLS) contain p-values for a Hausman test for the difference between OLS and 2SLS estimates (under the null hypothesis of OLS being unbiased and efficient, $(\gamma_{j,\text{OLS}} - \gamma_{j,\text{2SLS}})^2/(\sigma_{j,\text{2SLS}} - \sigma_{j,\text{OLS}})^2 \sim \chi^2_{(1)}$)

Table 4: Results based on pooled cross sections and retirement defined by end of last job

		Immediate Recall	**Delayed Recall**	**Total Recall**	**Serial 7's**	**Numeracy**	**Self-Rated Memory**
White Collar	OLS	-0.00269	-0.00307	-0.00580	-0.00093	-0.00094	-0.00167
	S.E.	0.00060	0.00073	0.00128	0.00047	0.00035	0.00038
	2SLS	-0.00074	0.00200	0.00075	0.00257	-0.00049	-0.00814
	S.E.	0.00450	0.00554	0.00961	0.00348	0.00255	0.00352
	p(OLS=2SLS)	0.617	0.291	0.432	0.245	0.836	0.039
Blue Collar	OLS	-0.00179	-0.00185	-0.00351	-0.00268	-0.00097	-0.00199
	S.E.	0.00044	0.00053	0.00092	0.00054	0.00030	0.00031
	2SLS	0.00739	-0.00050	0.00681	-0.00445	-0.00017	0.00341
	S.E.	0.00578	0.00558	0.01038	0.00704	0.00400	0.00439
	p(OLS=2SLS)	0.085	0.790	0.275	0.786	0.830	0.186
N	total	23013	22932	22932	22820	7751	23170
	clustered	6090	6083	6083	6070	4820	6110

Note: Regressions with pooled blue- and white-collar workers. The instruments, retirement status, and the quadratic age trend are interacted with the indicator for worker type. Further controls includes education, race, ethnicity, and wave dummies. Data are pooled cross-sections of all waves between 1996 and 2006. Heteroskedasticity-robust standard errors are adjusted for clustering. The rows labeled p(OLS=2SLS) contain p-values for a Hausman test for the difference between OLS and 2SLS estimates (under the null hypothesis of OLS being unbiased and efficient, $(\gamma_{j,\mathrm{OLS}} - \gamma_{j,\mathrm{2SLS}})^2/(\sigma_{j,\mathrm{2SLS}} - \sigma_{j,\mathrm{OLS}})^2 \sim \chi^2_{(1)}$)