Amin, Vikesh et al.

Working Paper

Does schooling improve cognitive abilities at older ages? Causal evidence from nonparametric bounds

Upjohn Institute Working Paper, No. 25-417

Provided in Cooperation with:
W. E. Upjohn Institute for Employment Research, Kalamazoo, Mich.

Suggested Citation: Amin, Vikesh et al. (2025) : Does schooling improve cognitive abilities at older ages? Causal evidence from nonparametric bounds, Upjohn Institute Working Paper, No. 25-417, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI, https://doi.org/10.17848/wp25-417

This Version is available at:
https://hdl.handle.net/10419/324812

6-6-2025

Does Schooling Improve Cognitive Abilities at Older Ages? Causal Evidence from Nonparametric Bounds

Vikesh Amin
Central Michigan University, amin1v@cmich.edu

Jere R. Behrman
University of Pennsylvania

Jason M. Fletcher
University of Wisconsin-Madison, IZA, and NBER

Carlos A. Flores
California Polytechnic State University, cflore32@calpoly.edu

Alfonso Flores-Lagunes
W.E. Upjohn Institute for Employment Research, IZA, and GLO, flores-lagunes@upjohn.org

See next page for additional authors

Upjohn Institute working paper ; 25-417

Does Schooling Improve Cognitive Abilities at Older Ages? Causal Evidence from Nonparametric Bounds

Authors

Vikesh Amin, *Central Michigan University*
Jere R. Behrman, *University of Pennsylvania*
Jason M. Fletcher, *University of Wisconsin-Madison, IZA, and NBER*
Carlos A. Flores, *California Polytechnic State University*
Alfonso Flores-Lagunes, *W.E. Upjohn Institute for Employment Research, IZA, and GLO*
Hans-Peter Kohler, *University of Pennsylvania*

Upjohn Author(s) ORCID Identifier


https://orcid.org/0000-0001-7996-7617

Published Version

In Demography 62(2): 515–541

W.E. UPJOHN INSTITUTE
FOR EMPLOYMENT RESEARCH

WORKING PAPER

Does Schooling Improve Cognitive Abilities at Older Ages? Causal Evidence from Nonparametric Bounds

Vikesh Amin
Central Michigan University

Jere R. Behrman
University of Pennsylvania

Jason M. Fletcher
University of Wisconsin-Madison, IZA and *NBER*

Carlos A. Flores
California Polytechnic State University

Alfonso Flores-Lagunes
W.E. Upjohn Institute for Employment Research, IZA and *GLO*
Email: Flores-Lagunes@upjohn.org

Hans-Peter Kohler
University of Pennsylvania

Upjohn Institute Working Paper 25-417

June 2025

ABSTRACT

We revisit much-investigated relationships between schooling and health, focusing on schooling impacts on cognitive abilities at older ages using the Harmonized Cognition Assessment Protocol in the Health & Retirement Study (HRS) and a bounding approach that requires relatively weak assumptions. Our estimated upper bounds on the population average effects indicate potentially large causal effects of increasing schooling from primary to secondary; yet, these upper bounds are smaller than many estimates from the literature on causal schooling impacts on cognition using compulsory-schooling laws. We also cannot rule out small and null effects at this margin. We do, however, find evidence for positive causal effects on cognition of increasing schooling from secondary to tertiary. We replicate findings from the HRS using older adults from the Midlife in United States Development Study Cognitive Project. We further explore possible mechanisms through which schooling may be working—such as health, SES, occupation and spousal schooling—finding suggestive evidence of effects through such mechanisms.

JEL Codes: I10, I26, J14

Key Words: Schooling, Cognition, Bounds, Aging, Partial-Identification

Acknowledgements: We thank Giuseppe Germinario for excellent research assistance.

1. Introduction

Does schooling have a causal effect on cognition at older ages? This is an important question for several reasons. First, cognitive health is a vital component of the overall health and well-being of older adults, who due to population aging represent an increasing segment of the population. In 2020, people aged 65 and above accounted for 17% of the US population, whereas this group only accounted for 5% of the population in 1920. The growth grate of the older age population over this period is almost five times that of the total population (Caplan 2023). Lower cognitive functioning is associated with a poorer quality of life due to difficulties in performing daily activities (Garrett et al. 2019), increased disabilities (Lee et al. 2005), and higher mortality risks (Batty et al. 2016). Understanding causal effects of schooling on older-age cognition can thus help determine whether investing in schooling will lead to health improvements for older adults.

Second, the causal effect of schooling on cognition of older adults has important implications for tackling Alzheimer's disease and related dementia (ADRD), a pressing public health issue. In 2023, 6.7 million Americans were living with ADRD, and this is projected to reach 13.8 million by 2060 (Alzheimer's Association 2023). The health care costs of ADRD are forecasted to increase from $345 billion in 2023 to $1 trillion in 2050 (Alzheimer's Association 2023). ADRD also imposes substantial costs to caregivers. In 2022, more than 11 million caregivers provided 17 billion hours of care worth $340 billion (Alzheimer's Association 2023). While age is the largest risk factor for ADRD, schooling has been identified as a modifiable early-life intervention to tackle ADRD. The 2024 Lancet Commission on Dementia Prevention, Intervention and Care estimated that 5% of dementia cases could be avoided by increased levels of basic schooling, and research on the US has found that increases in schooling attainment are significantly associated with declining dementia trends (Hudomiet et al. 2022).

One mechanism through which schooling can reduce dementia risks is by improving cognitive abilities at older ages. Neuropsychological testing is widely used to determine whether individuals meet criteria for dementia diagnoses. Higher schooling attainment is associated with better performance on cognition tests (Opdebeek et al. 2016). This entails a lower likelihood that an individual is diagnosed with dementia given existing thresholds, and thus a later age at dementia onset. Small delays in age at onset can lead to large reductions in population disease burdens. For example, Brookmeyer et al. (1998) projected that an intervention taken in 2015 to delay the onset of dementia by one year would reduce global cases in 2050 by 10% (12 million). Given the potentially large effect of schooling on overall dementia cases operating through the age of dementia onset, understanding whether schooling has a causal effect on cognition of aging people is vital for assessing future dementia trends, as cohorts reaching older ages will have higher levels of schooling as a result of schooling expansion in the 20th century.

Third, poor cognition and ADRD have important financial ramifications. Lower cognitive function at older ages is associated with less favorable financial outcomes. Banks & Oldfield (2007) found that low levels of numeracy were associated with lower levels of wealth and lower probability of owning a private pension in the English Longitudinal Study of Ageing (ELSA). In the US Health & Retirement Study (HRS), Angrisani & Lee (2019) found that declines in cognition scores of the household financial decision maker were associated with reductions in household wealth. Nicholas et al. (2021) found individuals with ADRD were more likely to miss bill payments six years before diagnosis and develop subpar credit scores 2.5 years before diagnosis. Thus, policies to increase schooling attainment may improve both the health and financial standing of older adults if schooling has a causal effect on cognition decades later in life at older ages.

Figure 1 illustrates mechanisms through which schooling could affect cognition at older ages. Schooling can improve cognition of older adults through reducing age-related brain pathology (brain maintenance) and by building cognitive reserve (Stern et al. 2023). It is hypothesized that schooling protects the brain from age-related pathology such as white matter damage which accumulates in all older adults, thus aiding brain maintenance. Cognitive reserve refers to the brain's ability to maintain successful cognitive performance despite age-related brain changes. Cognitive stimulation that occurs during schooling is thought to build cognitive reserve. Schooling attainment is a marker of cognitive reserve, and higher schooling attainment is hypothesized to allow individuals to process and store information that allows for normal cognitive functioning despite brain aging (Stern 2002). Some empirical evidence found that schooling was associated with cognitive reserve but not brain maintenance (Zahodne et al. 2019).[1]

There may be indirect effects of schooling through SES and health at middle and older ages. More schooling is associated with greater income, enabling the purchase of basic health-enhancing resources. More-schooled individuals may be more likely to work in cognitively stimulating occupations or more frequently engage in cognitively stimulating activities (e.g., social/cultural clubs) that build cognitive reserve. Similarly, more schooling is associated with better self-reported health and lower likelihoods of depression and smoking, which are associated with better cognition. Another indirect channel is marriage. Highly schooled individuals tend to also have highly schooled spouses, and spousal schooling is positively associated with older-age cognition (Liu et al. 2024; Saenz et al. 2020; Xu 2020). Spousal schooling can influence older-age

[1] Using two large samples of older adults without dementia at baseline, they found that schooling was not associated with white matter hyperintensities, suggesting no relationship between schooling and brain maintenance. They found evidence in support of cognitive reserve as memory scores of individuals with higher schooling attainment were less affected by increases in white matter hyperintensity volume compared to those with less schooling.

cognition by promoting healthier behaviors and increasing household SES (income/wealth) and cognitive reserve through engagement in cognitively stimulating activities.

Alternatively, associations between schooling and cognition at older ages may reflect the influence of earlier abilities or other unobserved confounders (e.g., genetics and family background) that are correlated with schooling and older-age cognition rather than causal relationships. Studies using schooling reforms as natural experiments have found evidence of a causal effect of schooling on cognition at ages 19–20 (Brinch & Galloway 2012; Lager et al. 2017; Xiao et al. 2017), suggesting that early-life cognition could be a confounding factor. Hence, it is important to determine whether the relationship between schooling and older age cognition is causal.

The literature on causal effects of schooling on cognition at older ages is surprisingly sparse. Some studies using changes in compulsory schooling laws to identify exogenous variation in schooling attainment (e.g., Glymour et al. 2008; Banks and Mazzonna 2012; Schneeweis et al. 2014; Gorman 2023) have found protective effects on immediate and delayed memory, with estimates showing an extra grade of schooling improved immediate and delayed memory scores by 0.08-0.50 standard deviations (SDs). However, findings for other cognitive domains are mixed.

Despite some advantages, estimates based on compulsory-schooling laws also have shortcomings. First, they capture the effects only for individuals whose schooling is causally increased by such laws, rather than effects for the general population. Second, they mostly represent effects of increasing schooling from primary to secondary and are not directly informative about effects of schooling at other parts of the educational distribution. Third, causal inference relies on the relatively strong assumption that school reforms affect cognition only

through their effect on schooling (the exclusion restriction), which may not always hold (e.g., Avendano et al. 2020).

We provide new evidence on the causal effect of schooling on cognition at older ages using the Harmonized Cognition Assessment Protocol in the HRS. Our contribution is to employ a nonparametric partial-identification approach (Manski and Pepper 2000), which provides bounds on the causal effect using relatively weak and arguably credible assumptions. Specifically, we assume (1) that there is positive selection into schooling such that individuals with higher schooling attainment have on average higher latent cognition and (2) more schooling does not worsen cognitive abilities. We then employ mother's schooling attainment as a monotone instrumental variable—a variable that is assumed to have a weakly increasing mean relationship with potential outcomes—to help tighten the bounds under these assumptions.

Our approach has several attractive features. *First*, it provides bounds on the population average treatment effect (ATE) as opposed to the average effect for a subpopulation, such as those for whom compulsory-schooling laws are binding. *Second*, it allows for arbitrary correlations between schooling and unobserved factors that can affect cognition. *Third*, it allows us to look at dose-response relations between schooling and cognition by providing bounds on the effect of increasing schooling at different parts of the educational distribution (e.g., going from being a high-school dropout to a high-school graduate, or from being a high-school to a college graduate). There may be important effects of obtaining credentials (high-school diploma; college degree) on cognition because credentials likely have large effects on mid-life conditions such as income and occupation. Nonlinear effects of schooling have been observed for other health outcomes, including mortality (Montez et al. 2012).

We find that there are potentially large effects of *completing* secondary schooling, with estimated bounds indicating that an extra grade of schooling *at most* increases immediate and delayed memory by 0.18 SDs when increasing schooling from primary to secondary. This estimated upper bound is smaller than many estimates from studies using compulsory-schooling laws. We also cannot rule out small and null effects at this margin. We obtain tighter bounds that indicate a statistically significant positive causal effect of increasing schooling from secondary to tertiary. An extra grade of schooling increases immediate and delayed memory scores by 0.03-0.10 SDs when transitioning from being a high-school to college graduate. Statistically positive effects are also found for several other cognition domains. We find suggestive evidence that this effect could work through better health at older ages, lower probabilities of being in poverty, higher probabilities of working in managerial and professional occupations (which likely involve cognitive stimulation) and having more-schooled spouses.

We view our estimates as providing important new and complementary evidence about the plausible magnitude of the causal effect under relatively weak assumptions. By using a completely different research design, our study contributes by triangulating evidence and increasing confidence that there exists a causal relationship between schooling and older-age cognition. Though we do not point-identify the causal effect, Mullahy et al. (2021), argue that partial-identification should be more prevalent in public-health and clinical research: rather than focusing on point estimates, base public-health recommendations, and policies on *ranges* of plausible effects. Finally, we replicate the findings from the HRS in a sample of older adults from the Midlife in United States Development Study Cognitive Project. This provides additional confidence in our results and highlights the value of using partial identification in different datasets to assess external validity.

2. Previous Studies

Most studies have used schooling variation arising from changes in compulsory-schooling laws within instrumental variable (IV) and fuzzy regression-discontinuity (RDD) designs.[2] Glymour et al. (2008) predicted grades of schooling in the 1980 US Census 5% sample using compulsory-schooling laws between 1907-1961. Predicted grades of schooling were then employed as an independent predictor of cognition in the HRS. Banks and Mazzonna (2012) used the ELSA and the 1947 increase in the minimum school leaving age from 14 to 15 with a fuzzy RDD. Both studies found large effects of schooling on memory. An extra grade of schooling was associated with a 0.34 (0.50) SD increase in memory in the HRS (ELSA). Glymour et al. (2008) found no effect of schooling on mental status. Banks and Mazzonna (2012) found that schooling improved executive functioning for men but not for women. Gorman (2023) employed the 1972 increase in the minimum school-leaving age from 15 to 16 in the UK and the Understanding Society dataset. She found that an additional grade of schooling increased memory scores by 0.42 SDs. She also found positive but imprecise effects on verbal fluency. Exploiting compulsory-schooling laws across Europe, Schneeweis et al. (2014) found an extra grade of schooling increased immediate (delayed) memory by 0.08 (0.09) SDs for older adults in Survey of Health, Ageing and Retirement in Europe. They found no causal effects of schooling on verbal fluency, numeracy, and orientation to-date.

A key assumption for causal inferences in these studies is the exclusion restriction—that school reforms affect cognition only through their effect on schooling—which could be violated in certain contexts. For example, using the 1972 schooling reform in the UK, Avendano et al.

[2] These studies are summarized in appendix table C1. Appendix Table C1 also shows two studies have used within-sibling comparisons. Herd and Sicinski (2022) found that more schooling is associated with higher memory scores for individuals in their 70s in Wisconsin. Fletcher et al. (2021) found higher fluid intelligence scores for more-schooled individuals in their 50s in the UK.

(2020) found that education did not improve mental health for individuals in their mid-50s, and that compulsory-schooling laws may affect later life mental health through channels other than increased schooling.[3] They argued and provided descriptive evidence that the reform forced young people who did not want to stay in school, but rather go to the labor market, to continue their education. These young people may have been negatively affected by being forced to stay in school in a stressful academic environment in which they were less likely to succeed compared to their peers. Courtin et al. (2019) also found that the increase in the minimum school leaving age from 14 to 16 in France 1959 increased depressive symptoms for women in their 60s. These findings imply that results in Gorman (2023) that are based on the 1972 reform, and more generally those that utilize compulsory-schooling laws for identification, may be biased because compulsory schooling laws possibly directly led to worse mental health, which in turn may have affected cognition (e.g., Donovan et al. 2017; Nafilyan et al. 2021). Hence, mental health is another channel through which compulsory-schooling laws could affect cognition later in life, which would violate the exclusion restriction. This assumption could also be violated if compulsory-schooling laws are correlated with school quality, which can affect cognition independently of level of schooling. Stephens and Yang (2014) showed that estimates of the effect of schooling on wages, unemployment, divorce, occupation in the US using compulsory-schooling laws as instruments became insignificant, and in many instances wrong-signed, when controlling for school quality.

In the presence of heterogenous effects, IV and fuzzy-RDD methods identify a local average treatment effect (LATE) for those individuals whose treatment is affected by the instrument ("compliers"). That is, these methods estimate the average effect of increasing schooling on cognition only for individuals who increased their schooling because of the

[3] They found that an extra grade of schooling increased the probability of having a mental health condition, and the probability of having depression/anxiety by about 30%.

compulsory-schooling laws (i.e., those whose schooling would have been lower in the absence of such laws). Previous IV and fuzzy-RDD studies are thus not directly informative about the population ATE, or about schooling effects at upper parts of the schooling distribution (e.g., college education), as compliers are generally individuals from the lower part of the schooling distribution (Clark and Royer 2013).[4]

In sum, the current evidence suggests that there likely is a causal relationship between schooling and memory, but findings for other cognition domains are mixed. Point identification of causal effects, though, rests on strong assumptions; the effects identified pertain to specific subpopulations and usually lower parts of the schooling distribution. In contrast, our approach employs relatively weak assumptions to provide bounds on the population ATE of increasing schooling at different parts of the schooling distribution.

3. Econometric Framework

Let $Y_i\ (t_1)$ and $Y_i(t_2)$ be two potential outcomes, i.e., the values of the outcome (older-age cognition) that individual i would obtain as a function of two different treatment or schooling levels (e.g. high-school graduation t_2 and less-than high-school t_1). We are interested in the population ATE of increasing schooling attainment from less-than high-school (t_1) to high-school graduation (t_2) on cognition test scores, defined as:

$$(1)\ \ \Delta\ (t_1, t_2) = E[Y(t_2)] - E[Y(t_1)]$$

[4] Exploiting variation in college availability and student loan regulations in Germany, Kamhöfer et al. (2019) found positive effects of college graduation on reading speed, reading comprehension and mathematical literacy for individuals in their early 50s.

Estimation of the ATE is complicated because the potential outcome $Y(t_2)$ is unobserved for individuals with schooling level different from t_2, and $Y(t_1)$ is unobserved for individuals with schooling level different from t_1. Letting S denote realized schooling, and using the law of iterated expectations to write the expected potential outcome $E[Y(t_2)]$ as:

$$(2)\ E[Y(t_2)] = E[Y(t_2)|S < t_2] * P(S < t_2) + E[Y(t_2)|S = t_2] * P(S = t_2) + E[Y(t_2)|S > t_2] * P(S > t_2)$$

The data identify the sample analogues of all the right-side quantities except of the counterfactuals $E[Y(t_2)|S < t_2]$ and $E[Y(t_2)|S > t_2]$, i.e., the average cognition under high school graduation (t_2) for individuals with realized schooling (S), respectively, less and higher than high school. A similar equation applies to $E[Y(t_1)]$. The bounding approach we employ consists of making assumptions to bound each one of the counterfactuals in the expressions for $E[Y(t_2)]$ and $E[Y(t_1)]$ to then bound the ATE $\Delta\ (t_1, t_2)$. The assumptions are outlined below and illustrated in Figure 2 with further details in appendix A.

Assumption 1: Bounded Support. This assumption exploits that the measures of cognition employed have a minimum (Y_{min}) and a maximum (Y_{max}) value, which are used in place of the counterfactuals (Figure 2 Panel A) to obtain a lower and an upper bound on each $E[Y(t_2)]$ and $E[Y(t_1)]$.

Assumption 2: Monotone Treatment Selection (MTS): MTS states that individuals with higher schooling attainment on average have weakly higher potential outcomes at every schooling level t. For example, when comparing high-school graduates to high-school dropouts, the MTS

assumption requires that average potential cognition score at any schooling level (e.g., less-than high-school, some college, college graduation) of high-school graduates is higher than that of high-school dropouts.

While the MTS assumption is untestable (since counterfactual outcomes are unobserved), it is plausible in our application. Economic models of human capital posit that individuals with higher innate ability have more schooling (Ben-Porath 1967), and polygenic scores for education and cognition (which can be interpreted as measures of innate ability) predict cognition at older ages (Fletcher et al. 2021; Herd and Sicinski 2022), indicating that higher innate ability is likely related to better cognition at older ages. Given that individuals with higher innate ability are more likely to have more schooling and better cognition at older ages, it is plausible that, on average, individuals with higher schooling attainment have higher potential cognition at all schooling levels.[5] More generally, MTS captures the notion that, relative to individuals who self-select into lower schooling levels, individuals who self-select into higher schooling levels are more likely to have pre-treatment characteristics that also make them more likely to have better average potential older-age cognition at any given schooling level, for instance, due to (on average) higher innate ability, better health inputs and better family background.

Figure 2 Panel B shows that under the MTS assumption, the observed mean cognition for those with schooling t_2 ($E[Y(t_2)|S = t_2]$)—e.g., high school—can be used as an upper bound (respectively, lower bound) for the mean potential cognition under high school for those with realized schooling less (higher) than high school, as $E[Y(t_2)|S < t_2] \leq E[Y(t_2)|S = t_2] \leq E[Y(t_2)|S > t_2]$.

[5] Note that correlations of unobserved factors (e.g., innate ability) with schooling and cognition are consistent with the MTS. What MTS rules out is the possibility that third factors affect cognition in such a way that, on average, at a given level of schooling, more-schooled individuals have worse potential cognitive performance than less-schooled individuals.

Assumption 3: Monotone Treatment Response (MTR): MTR assumes that more schooling does not decrease cognitive ability at older ages for any individual; that is, for two schooling levels t_1 and t_2 with $t_2 > t_1$, $Y_i(t_2) \geq Y_i(t_1)$. MTR is thus an assumption about the (weak) ranking of the potential outcomes for the same individual—MTR compares potential cognition under high-school graduation versus the potential cognition under less-than high-school graduation for the same individual. This is different from the MTS assumption, which compares means of the same potential outcome (e.g., cognition under high-school graduation) for two different subpopulations defined by their observed levels of schooling (e.g., high-school graduates versus high-school dropouts).

The MTR assumption is stronger than the MTS assumption as it is required to hold for each individual, rather than on average. The MTR assumption would be violated if more schooling leads to worse cognitive performance for some individuals. One could argue that more schooling may worsen mental health for individuals who work in stressful jobs, or for individuals forced to stay in school by compulsory-schooling laws, and that such deterioration in mental health could lead to worse cognition. The MTR assumption does not rule out such negative channels, but rather assumes that the positive channels linking schooling to cognition dominate; for example, the negative impacts of poor mental health on cognition are outweighed by the positive impacts of schooling on brain maintenance, cognitive reserve, assortative mating, and mid-life health and SES conditions.

There are theoretical models to suggest that the MTR assumption holds. In the Grossman (1972) model of health production, schooling directly increases health production by increasing the marginal productivity of health inputs or behaviors (productive efficiency) and by enhancing

individuals' abilities to acquire and process health information (allocative efficiency). Cognition at older ages is a component of overall health, which is increased through schooling's effect on productive and allocative efficiency. The MTR assumption is also consistent with theories of cognitive reserve (Stern 2002).[6]

Figure 2 Panel C illustrates how MTR tightens bounds on $E[Y(t_2)]$. Intuitively, for the counterfactual $E[Y(t_2)|S = t_1]$ (e.g., mean cognition under high-school for those with no high school), MTR provides the lower bound $E[Y(t_1)|S = t_1]$ (mean observed cognition for those with no high school), since more schooling cannot decrease potential cognition, implying $E[Y(t_2)|S = t_1] \geq E[Y(t_1)|S = t_1]$. Panel D shows how tighter bounds on $E[Y(t_2)]$ can be obtained by combing the MTS and MTR assumptions. Manski & Pepper (2000) show that a testable implication of the MTS+MTR assumptions is that observed mean cognition scores are weakly increasing in schooling attainment. This testable implication will fail if the assumptions are not satisfied in the data, providing a check on them.

Assumption 4: Monotone Instrumental Variable (MIV): Each of the bounds under the previous assumptions can be narrowed by using a MIV, which is a variable that has a monotone (weakly increasing or weakly decreasing) mean relationship with the potential outcomes $Y(t)$. We use mothers' schooling attainments as the MIV, thus assuming individuals' mean potential cognitions at older ages are weakly increasing (i.e., not strictly decreasing) in their mothers' schooling levels.

[6] The active cognitive-reserve hypothesis posits that individuals with more education make more efficient use of brain networks and process tasks more efficiently, leading them to experience less cognitive decline from brain aging compared to less-educated individuals. The common cause hypothesis argues that if cognition declines in age come from a common cause, then the cognition of higher educated individuals will decline at a similar rate to the population rate. However, more educated individuals will continue to perform at a higher level at a given age because of greater baseline brain reserve. The compensation hypothesis states that education allows more cognitive domains to fully develop, and once brain aging affects cognition, the domains not affected compensate for declines in the other cognitive domains.

Mother's schooling attainment is a natural MIV as studies have shown that higher parental schooling, and more generally childhood SES, are associated with better cognition at older ages (see Greenfield and Moorman 2019 for a review).

The MIV assumption is weaker than the exclusion restriction that would require that mothers' schooling affects children's older-age cognition only through effects on children's schooling. The exclusion restriction is unlikely to be satisfied because mothers' schooling can affect children's older-age cognition through many other channels. Mothers with more schooling may have children with higher innate abilities, who in turn obtain higher schooling levels and have better older-age cognitions. More-schooled mothers may also provide better nutrition to their children or more cognitively stimulating home environments that affect neurocognitive development (Hackman and Farah 2009). These channels are allowed by the MIV assumption and help to justify its plausibility. This assumption does not require the MIV (mothers' schooling) to have causal effects on outcomes (children's older-age cognitions).[7]

For mothers' schooling, the MIV assumption implies that if we take two subsamples where individuals counterfactually have the same level of schooling but two different levels of mothers' schooling (e.g., mothers are high-school dropouts and mothers are high-school graduates), mean potential cognition $E[Y(t)]$ will be weakly higher for the sample whose mothers are high-school graduates. Note that this inequality holds weakly, so it also allows for such means to be equal. The inequality is also assumed to hold for all schooling levels t, and it does not need to hold for every individual in the two subsamples, as the MIV assumption refers to average (rather than individual)

[7] Mothers' schooling is positively associated with characteristics such as children's cognition (Carneiro et al. 2013; Cave et al. 2022; Dickinson et al. 2016; Magnuson 2007) and their schooling attainments (De Hann 2011; Holmlund et al. 2011; Sacerdote et al. 2002), which in turn affect older-age cognitions.

potential outcomes. Lastly, the MIV assumption is untestable as counterfactual outcomes are unobserved.

To illustrate how the MIV narrows bounds, consider using it along with the MTS+MTR bounds. We can divide the sample into bins defined by the values of mothers' schooling and compute the MTS+MTR bounds within each bin. The MTS+MTR+MIV bounds on $E[Y(t_2)]$ are then obtained by taking the weighted average over all the conditional-on-MIV bounds, resulting in bounds that are (weakly) narrower than without using the MIV.

Bounds on the ATE $\Delta\,(t_1, t_2)$: Let $LB^A\ E[Y(t)]$ and $UB^A\ E[Y(t)]$ denote, respectively, the lower and upper bounds for the mean potential outcome at schooling level t under the different assumptions A={NA, MTS, MTR, MTS+MTR, MTS+MTR+MIV}, where NA denotes the "no-assumptions" bounds using only Assumption 1. Bounds on the ATE of increasing schooling from t_1 to t_2 for a given set of assumptions are calculated by:

$$(3)\ LB^A E[Y(t_2)] - UB^A E[Y(t_1)] \leq \Delta\,(t_1, t_2) \leq UB^A E[Y(t_2)] - LB^A E[Y(t_1)]$$

Bounds on the ATE for other schooling margins (e.g. increasing schooling from high-school graduation to some college, or from high-school to college graduation) are computed analogously. Note that under the MTR assumption the lower bound on $\Delta\,(t_1, t_2)$ is never below zero, because the MTR rules out the possibility that more schooling worsens cognitive abilities.

Estimation and Inference: All bounds are estimated by plugging in sample analogs for the expectations and probabilities in the corresponding bounds' expressions provided in appendix A.

Inference is undertaken by constructing confidence intervals that cover the true value of the population ATE with a specified probability. Estimation and inference for bounds using the MIV assumption is non-standard, requiring dealing with biases arising from taking min and max (intersections) over several candidate bounds. We deal with these issues by applying the methodology of Chernozhukov et al. (2013) to obtain all the estimated bounds and confidence intervals, including those not using the MIV assumption. Section A of the online appendix provides further details. We implement the methods using the user-written STATA command *mpclr* (Germinario et al. 2021).

4. Data

The HRS is a US nationally representative longitudinal survey of individuals over age 50 and their spouses that started in 1992. The initial HRS cohort consisted of persons born in 1931–41 and their spouses of any age. A second study, Asset and Health Dynamics Among the Oldest, was fielded the next year to capture an older birth cohort, those born in 1890–1923. In 1998, the two studies were merged, and, to make the sample fully representative of the older US population, two new cohorts were enrolled, the Children of the Depression, born in 1924–1930, and the War Babies, born in 1942–1947. The HRS now employs a steady-state design, replenishing the sample every six years with younger cohorts to continue making it fully representative of the population over age 50.

4.1 Schooling Attainment

We obtain the schooling attainment of HRS participants and their mothers from the RAND HRS dataset (version V1), which is measured with grades of schooling ranging from 0–17. We

discretize participants' grades of schooling into high-school dropout (<12), high-school graduation (12), some college education (13-15) and college graduation (≥16). Mothers' schooling is discretized into high-school dropout, high-school graduate and more than high school.

4.2 Cognition Measures

Cognition measures come from the Harmonized Cognition Protocol Assessment (HCAP), which was initiated in 2016. Participants were selected to be part of HCAP if they were 65 years or older and had completed the 2016 interview. Though the HRS includes spouses/couples, there are no spousal/couple pairs in HCAP. Of those eligible for HCAP, the HRS randomly selected half of uncoupled respondents, and one respondent from each coupled household.

HCAP consisted of two parts, a respondent interview and an informant interview. In the respondent interview, participants completed comprehensive, in-person neuropsychological assessments (see appendix B for test descriptions) that took about one hour. Immediately afterwards an individual nominated by each HRS respondent completed an informant interview in another room answering questions on the respondent's functioning and changes in abilities over the last 10 years. Of the eligible 4,425 participants, 3,496 completed the HCAP interview. There were 149 cases where the HRS respondents were not able to conduct interviews and only the informant interviews were conducted.

Appendix Table C2 shows the order in which the cognitive tests were undertaken, cognitive domains assessed, numbers of observations and numbers of missing observations. For most cognition tests, the number of missing observations is low (<50; 2% of HCAP respondents). For some tests, the number of missing observations is higher (e.g., letter cancellation has 150

observations imputed). The HRS has imputed missing cognition scores, so we do not lose data due to missing cognition scores in our analysis.

To aid comparisons with the literature, we primarily focus on composite scores for immediate and delayed memories constructed by summing up the scores on the HRS Telephone Interview for Cognitive Status (TICS), three trials of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) word list-immediate, three trials of the CERAD word list-delayed, CERAD constructional praxis-delayed, story-recall immediate, and story-recall delayed tests. The other cognition outcomes we examine are: Mini Mental State Examination (MMSE, which is seen as a measure of global cognition), recognition memory (sum of scores on CERAD word list-recognition and story-recall recognition), verbal fluency (animal naming test), executive function (Raven's progressive matrices test), attention/speed (sum of scores on backward counting and letter cancellation tests), and visuospatial (CERAD constructional praxis-immediate).

5. Results

5.1 Descriptive Statistics

Summary statistics for our analytical sample are shown in Table 1 column 1. We have an analytical sample of 3,072. From the full HCAP sample of 3,496 observations, we lose 149 observations where no respondent interview was conducted. We then drop respondents with missing data on schooling (4 observations), mothers' schooling (269 observations), and race (2 observations). The average age is 76 years (range of 65–102 years), 61% are female, and 73% are non-Hispanic white. Mothers' schooling is concentrated at the lower end of the schooling distribution. Over half (54%) of individuals had mothers who never graduated from high school, and 32% of individuals had mothers who graduated from high school. Only 14% of individuals

had mothers who had more than high-school education. Average grade of schooling for HCAP participants is 12.92, with 18% of individuals being high-school dropouts and 25% being college graduates.

Columns 2-5 provide summary statistics by gender and race. Although men have higher schooling attainment than women, they do not outperform women on all cognition domains. Women score higher on average on the MMSE, immediate and delayed memory, and recognition memory. On average, men perform better than women on verbal fluency, executive function, attention/speed, and visuospatial tests. Non-Hispanic white and non-Hispanic black individuals have higher schooling attainment than Hispanic individuals. Average cognition scores on all domains are higher for non-Hispanic white individuals compared to non-Hispanic black and Hispanic individuals.

Average cognition scores by schooling attainment are in appendix Table C3. Cognition scores are increasing in schooling attainments, which is consistent with the testable implication of the MTS+MTR assumptions. Finally, at equivalent levels of schooling, the average score on our composite measure of immediate and delayed memory is higher for women than men. For example, the average immediate and delayed memory score for women (respectively, men) who are high-school dropouts is 38.98 (34.80), and for those who are college graduates is 58.96 (53.71). These findings are consistent with descriptive evidence from Angrisani et al. (2020) showing that the schooling gap by gender is negligible on cognition tests that do not require numeracy or literacy.

5.2 Results for Immediate and Delayed Memory

Panel A of Table 2 displays the results for increasing schooling from high-school dropout to high-school graduate. The OLS estimate (column 2) indicates that the memory score of high-school graduates is on average 9.3 points higher than that of high-school dropouts. This represents an average effect of 0.58 SDs relative to the SD of high-school dropouts (the control group). Columns 3–7 show the estimated bounds and 95% confidence intervals under different sets of assumptions. The NA bounds are wide, indicating that the true average causal effect of completing secondary schooling could at worst lower memory scores by 71.80 points and at most improve scores by 74.44 points. Adding the MTS assumption—that individuals with higher schooling attainment have on average higher potential cognition—substantially reduces the estimated upper bound. Completing secondary schooling under the MTS assumption (column 4) at most increases memory scores by 34.47 points. The MTS bounds are still wide and include zero. Adding the MTR assumption by itself (column 5) restricts the lower bound mechanically to zero, because MTR rules out the possibility that schooling worsens cognition. The combination of the MTS and MTR assumptions (column 6) provides considerably tighter bounds compared to previous bounds. Completing secondary schooling at worst has no effect and at most increases the memory score by 12.92 points. This represents an effect of at most 0.81 SDs relative to the control SD. To tighten the MTS+MTR bounds we use mothers' schooling as a MIV with three bins (high-school dropout, high-school graduate, more than high school). Adding the MIV to the MTS+MTR assumptions slightly reduces the estimated upper bound to 11.48 points (0.72 SDs) in column 7.[8]

How does the range of causal effects for completing secondary schooling under the MTS+MTR+MIV assumptions compare with IV and fuzzy-RDD estimates from studies using

[8] Adding the MIV to the MTS (MTR) assumption also only leads to slight tightening of bounds. The ATE of going from being a high-school dropout to high-school graduate under the MTS+MIV (MTR+MIV) assumptions is

compulsory-schooling laws? First note that our bounds are on the population ATE of completing secondary schooling and are not directly comparable to prior IV and fuzzy-RDD estimates. While these latter estimates also capture effects at the lower part of the schooling distribution, they estimate a LATE for individuals who continued their schooling because of compulsory-schooling laws (the compliers), which may differ from the effect for individuals who would have remained in school regardless of the compulsory-schooling laws (Clark and Royer 2013), and thus from the population ATE. The bounds indicate that there is potentially a substantial effect of completing secondary schooling—0.72 SDs. This implies that an additional grade of schooling increases memory scores by at most 0.18 SDs, given a four-grade difference in schooling between high-school graduates and dropouts. Our estimated upper bound is thus substantially smaller than estimates of an extra grade of schooling for the US (0.34 SDs; Glymour et al. 2008) and UK (0.42–0.50 SD; Banks and Mazzonna 2012; Gorman 2023), but possibly larger than the IV estimates identified for Europe (0.08–0.09 SDs; Schneeweis et al. 2014). The fact that the IV estimate for the US in Glymour et al. (2008) is above our estimated upper bound may be interpreted as reflecting treatment-effect heterogeneity. Since compulsory-schooling laws are most likely to affect the schooling levels of individuals who would otherwise have relatively low schooling (Card, 1999), the average effect for the compliers being larger than for the population would be consistent with these individuals having higher marginal returns to additional years of secondary schooling in terms of cognition at older ages relative to the overall population. A similar reasoning has been used before in the context of estimating the effect of schooling on earnings, where Card (1999) points to possible differences in the returns to education as a potentially important reason why IV estimates of this effect based on compulsory-schooling laws tend to exceed corresponding

between −48.86 and 33.01 (0 and 52.82). MTS+MIV and MTR+MIV bounds for all schooling margins are shown in appendix Table C4.

OLS estimates. Finally, note that our bounds include the OLS estimates, whereas prior IV estimates are larger than OLS estimates.[9]

Panel B shows results for increasing schooling from high-school graduation to some-college education. Under the MTS+MTR+MIV assumptions going from high school to some-college education increases memory scores by 0–4.10 points (0–0.24 SDs). The MTS+MTR+MIV bounds in panel C show that going from having some college education to graduating from college increases memory scores by 0–5.26 points (0.32 SDs). We note that the estimated upper bounds under the MTS+MTR+MIV assumptions at both of these schooling margins are much smaller compared to those from completing high school.

Panel D provides results for increasing schooling from high-school to college graduation. Here, we obtain fairly tight bounds under the MTS+MTR+MIV assumptions. Increasing schooling from secondary to tertiary increases average memory score by 1.87–6.74 points (0.11–0.39 SDs).[10] The estimated bounds exclude zero and the OLS estimate (9.64), as does the 95% confidence interval, which implies that the true effect is between 1.14 and 7.75 points (0.07–0.45 SDs). Given a difference of four grades of schooling between college and high-school graduates, the estimated bounds (95% confidence interval) imply that an additional grade of schooling increases memory scores by 0.03–0.10 (0.02–0.11) SDs.

We also obtain informative bounds under the MTS+MTR+MIV assumptions for increasing schooling from primary to tertiary in panel E. The bounds indicate that the average causal effect

[9] With heterogeneous treatment effects, OLS estimates the population ATE while IV estimates a LATE. Differences between OLS and IV estimates may come from possible bias in OLS estimates or from the fact that OLS and IV methods estimate effects for different populations.

[10] Appendix Table C5 shows estimated bounds on the mean potential memory scores for each schooling level t ($E[Y(t)]$). For example, following equation (3) and under the MTS+MTR+MIV assumptions, the lower bound on the average memory-score effect from secondary (t_2) to tertiary (t_4) in Table 2 is positive because the estimated lower bound on E[Y(t_4)] (52.35) is greater than the estimated upper bound on E[Y(t_2)] (50.51). See note to online appendix Table C5 for further details.

is between 3.43–16.30 points (0.22–1.02 SDs). The 95% confidence interval excludes the OLS estimate (18.91) and zero, implying a statistically significant average memory effect of at least 0.22 SDs of increasing schooling from primary to tertiary.

Figure 3 gives MTS+MTR+MIV bounds by gender and race at the different schooling margins. Across both gender and race, there are possibly large effects of completing high school with estimated upper bounds of 0.77 SDs for men, 0.80 SDs for women, 0.65 SDs for non-Hispanic white individuals, 0.63 SDs for non-Hispanic black individuals and 0.90 SDs for Hispanic individuals. These bounds are also consistent with moderate, small, and null effects. In general, it is difficult to draw strong conclusions regarding effect heterogeneity because the bounds overlap and thus there are no statistically significant differences by gender and race. When examining differences between college and high-school graduates non-Hispanic white individuals have the narrowest bounds (0.09–0.34 SDs). In contrast, non-Hispanic black and Hispanic individuals have wider bounds, and the confidence intervals include zero. Finally, increasing schooling from primary to tertiary statistically significantly increases average memory scores across all genders and races.

5.3 Results for Other Cognition Domains

We focus on results from increasing schooling from primary to secondary (for comparisons with studies using compulsory-schooling laws) and from secondary to tertiary (to see if the informative bounds for memory replicate).[11] Figure 4 Panel A presents results for completing secondary schooling, indicating that there could be zero, small, or potentially large causal effects. The literature has found mixed results for cognition domains other than memory, and given the

[11] See appendix Table C6 for full results.

width of the bounds, we cannot make strong comparisons with previous studies. For example, the estimated upper bounds indicate that completing secondary schooling increases verbal-fluency scores at most by 0.67 SDs, implying that an additional grade of schooling increases verbal fluency by at most 0.17 SDs. This is larger than the IV estimate in Gorman (2023), who found that an extra grade of schooling increased verbal fluency by 0.05 SDs but was imprecisely estimated. The bounds do exclude IV estimates for verbal fluency in Schneeweis et al. (2014), which were all negative and statistically insignificant.

Panel B shows results for increasing schooling from secondary to tertiary. All the bounds statistically exclude zero (marginally for recognition memory and visuospatial), implying statistically significant average effects of increasing schooling from secondary to tertiary on these cognitive measures. The tightest bounds are obtained for the MMSE. Transitioning from being a high-school to college graduate increases MMSE scores by 0.08–0.27 SDs. The width of the bounds is similar for verbal fluency, executive function, attention/speed, indicating average causal effects of about 0.05–0.40 SDs. All the bounds and 95% confidence intervals (except for MMSE and recognition memory) exclude the OLS estimates. These results highlight the potential role that increasing schooling from secondary to tertiary can have in improving cognitive abilities at older ages.

Gender and race specific bounds under the MTS+MTR+MIV assumptions are shown in Figure 5. While there are no statistically significant differences, there is some suggestive evidence of racial differences for attention/speed, where we can only statistically rule out null effects for non-Hispanic black individuals (with bounds indicating effects between 0.26–0.47 SDs). Effects of increasing schooling at other parts of the schooling distribution by gender and race are shown

in appendix Figures C1-C4. As before, there are no statistically significant gender or race differences.

5.4 Possible Mechanisms

This section investigates effects of schooling on possible pathways through which schooling can affect older-age cognition. Panel A in Table 3 presents MTS+MTR+MIV bounds on effects of increasing schooling for HCAP respondents on the probability of reporting poor/fair health, body-mass index (BMI), depressive symptoms measured by the Center for Epidemiologic Studies Depression Scale (CES-D; scale 0-8), probability of reporting ever smoked, probability of reporting ever diagnosed with high blood pressure and probability of reporting not doing vigorous exercise in the 2016 HRS survey. Since schooling is negatively correlated with these measures, we employ non-positive versions of the MTS and MTR assumptions and assume that mothers' schooling has a weakly decreasing relationship with the mean potential outcomes of these measures.[12] Estimated bounds show that college graduates are 2.1–11.4 percentage points less likely to be in poor/fair health and 2.1–5.9 percentage points less likely to have ever smoked compared to high-school graduates. The CES-D score of college graduates is also 0.012–0.34 points lower than high-school graduates. While the estimated bounds exclude null effects, the 95% confidence intervals do not. The results thus provide some suggestive evidence that the effects of increasing schooling from secondary to tertiary on older-age cognition could operate partly through better health at older ages.

[12] Taking depressive symptoms as an example, the MTR assumption now requires that more schooling does not increase depressive symptoms, and the MTS assumption states that individuals with higher schooling attainment do not have strictly higher mean potential depressive symptoms at every schooling level. The MIV assumption states that the individuals' mean potential depressive symptoms are not strictly increasing in their mothers' schooling levels.

Panel B column 1 reports effects of schooling on the probability of the respondent's household being in poverty the last calendar year. Going from being a high-school to college graduate reduces the probability of being in poverty by 1.3–5.7 percentage points and the 95% confidence interval marginally excludes zero. We use information on the occupation code for the job with the longest tenure from the RAND HRS 2016 wave to create a dummy variable for having worked in managerial/professional occupations the longest. This may be an important channel because managerial/professional occupations are likely to be cognitively stimulating, which enhances cognitive reserve and protects against cognitive decline. Increasing schooling from secondary to tertiary statistically significantly increases the probability of having longest tenure in managerial/professional occupations by 2.2–9 percentage points (column 2), and the probability of having spouses who are college graduate by 6.2–28.4 percentage points (column 3). Overall, these results suggest that the effects of increasing schooling from secondary to tertiary on cognition could work through the channels in Panel B.

5.5 Robustness Checks

We conducted two robustness checks. First, our main analysis uses grades of schooling to group respondents and their mothers into educational groups. We investigated whether bounds would be narrower when using respondents' grades of schooling as the treatment and mother's grades of schooling (with three bins) as the MIV. Results are shown in appendix Table C7. Our main findings on the effects of increasing schooling from high-school to college graduation are robust to this alternative coding scheme. We find that increasing schooling from 12 to 16 grades increases immediate and delayed memory scores by 0.14–0.27 SDs, memory recognition by 0.04–0.17 SDs, MMSE by 0.10–0.23 SDs, verbal fluency by 0.10–0.34 SDs, executive function by 0.25–

0.33 SDs, attention/speed by 0.24–0.30 SDs, and visuospatial by 0.10–0.34 SDs. With this coding scheme we cannot statistically exclude null effects for MMSE, recognition memory, verbal fluency, and visuospatial, whereas in Figure 4 all the 95% confidence intervals exclude zero (marginally for recognition memory and visuospatial).

Second, we examined the robustness of the estimated bounds to attrition in the HRS through inverse probability weighting. We first performed a logit regression on the probability of being in the HRS 2016 survey as a function of year of birth, gender, schooling, mother's schooling, self-reported health, mental health (CES-D score), cognition scores, and BMI.[13] Self-reported health, BMI, CES-D, and cognition scores were averaged across the first observed wave through the last observed wave (not including 2016). Covariates with missing values (mainly mother's schooling) were imputed with their sample mean and controlled for in regressions with missingness dummy variables. The inverse of the predicted probabilities was then used to weight the observations when computing the OLS estimates and estimated bounds. Results are given in appendix Table C8, which are similar to our main findings.

5.6 Replication in the Midlife in United States Development Study (MIDUS)

For external validity, we examined schooling effects on cognition for a sample of older adults in the MIDUS, which is a national sample of 7,108 adults aged 25–74 first interviewed in 1995–96. Nine years later the second wave (MIDUS 2) included data from about 75% (N=4,963) of the original respondents. We use the MIDUS 2 Cognitive Project where 4,512 participants undertook the Brief Test of Adult Cognition by Telephone (BTACT). The BTACT included

[13] For cognition, we used the summary measure of cognition in the RAND dataset, which sums up scores from (1) a 10-word immediate and delayed recall tests of memory; (2) a serial 7s subtraction test of working memory; (3) counting backwards to assess attention and processing speed; (4) an object-naming test to assess language; and (5) recall of the date and president and vice-president to assess orientation.

measures of memory (immediate and delayed recall of 15 words), inductive reasoning (number series; completing a pattern in a series of 5 numbers), verbal fluency (the number of words produced from the category of animals in 60 seconds, as in the HRS), processing speed (backward counting, as in the HRS) and working memory (backward digit span; the highest span achieved in repeating strings of digits in reverse order). Despite its brief length, the BTACT is a reliable and valid measure of cognition (Lachman et al. 2014).

We restricted our analysis to 1,016 individuals aged 65 years or older with data on schooling and mothers' schooling. Summary statistics are in appendix Table C9. The average age and proportion of women in the MIDUS is similar to the HRS, as well as the distribution of mothers' schooling. Average grades of schooling are higher in the MIDUS (13.79) than in the HRS (12.92), reflecting the higher proportion of white individuals in the MIDUS (95% vs 73%). Average grades of schooling of white individuals in our HRS analytical sample (13.49) is similar to the MIDUS.

Appendix Figure C5 presents a comparison of results for increasing schooling from (a) primary to secondary and (b) secondary to tertiary (full results are in appendix Table C10). Similar to the HRS, the MIDUS bounds show that the average causal effect of completing secondary schooling could be zero, small, equal to OLS estimates, or potentially larger (but at most about 0.50–0.55 SDs, depending on the specific measure). The OLS estimates and the width of the bounds are quite similar, especially for immediate and delayed memory. In the MIDUS, estimated bounds for increasing schooling from secondary to tertiary on immediate and delayed memory are tight (0.10–0.17 SDs), but the 95% confidence interval is much wider (0–0.32 SDs). Similarly, the

bounds for verbal fluency and attention/speed do not statistically exclude zero, whereas they do in the HRS. In both datasets, all the estimated upper bounds are smaller than the OLS estimates.[14]

6. Summary

Does schooling have a causal effect on cognition at older ages? The evidence for this important question is surprisingly limited, given growing ADRD cases, the recognition of schooling as the largest non-biological life-cycle intervention for ADRD, and the many associations without attempts to provide casual estimates between schooling and various dimensions of aging. We contribute to the literature by employing a partial-identification approach to determine a range of plausible values for the population average causal effect of schooling on cognition in the HRS, under weak assumptions. We find that the average causal effect of increasing schooling from primary to secondary levels on immediate and delayed memory could be zero, small, or potentially large, but no more than 0.72 SDs. The estimated upper bound implies that an additional grade of schooling increases memory scores by at most 0.18 SDs. This is substantially smaller than estimates from studies using compulsory-schooling laws for identification, where estimates represent a LATE only for those who increase their schooling due to these laws. We also reach similar conclusions for global cognition, verbal fluency, executive function, recognition memory and visuospatial. We further provide new evidence that there are important effects of schooling on older-age cognition at other parts of the schooling distribution. This is critical because the previous literature using compulsory-schooling laws for identification obtaining LATE estimates for secondary-school completion, and the effects may differ at other points of the schooling distribution. We obtain a fairly narrow range of estimated average causal effects of

[14] We also estimated bounds on the effects of schooling on cognition for individuals aged 25–50 years, finding that increasing schooling from secondary to tertiary increases memory scores, verbal fluency, and attention/speed, respectively, by 0.03–0.18, 0.06–0.39, and 0.05–0.71 SDs, statistically ruling out null effects for the last two. Full results are in appendix Table C11.

increasing schooling from secondary to tertiary on all cognition domains. Moreover, these estimated bounds statistically rule-out zero effects for all the cognition domains (marginally for recognition memory and visuospatial). For example, an extra grade of schooling increases average immediate and delayed memory by 0.03–0.10 SDs when transitioning from being a high-school to a college graduate. Finally, we find some suggestive evidence that these effects may work through having higher probabilities of working in cognitively stimulating (managerial/professional) occupations, more-schooled spouses, higher SES and better health at older ages. Thus, our analyses lead to a more nuanced and extended understanding of the impacts of different levels of schooling on cognition at older ages in the US.

References

Alzheimer's Association (2023). 2023 Alzheimer's Disease Facts and Figures. Alzheimer's Dementia 19 (4), 1598-1695.

Angrisani, M., & Lee, J. (2019). Cognitive decline and household financial decisions at older ages. *Journal of the Economics of Ageing*, *13*, 86-101.

Angrisani, M., Jain, U., & Lee, J. (2020). Sex differences in cognitive health among older adults in India. *Journal of the American Geriatrics Society*, *68*(Suppl. 3) , S20-S28.

Avendano, M., de Coulon, A., & Nafilyan, V. (2020). Does longer compulsory schooling affect mental health? Evidence from a British reform. *Journal of Public Economics*, *183*, 104137. https://doi.org/10.1016/j.jpubeco.2020.104137

Banks, J., & Mazzonna, F. (2012). The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. *Economic Journal*, *122*(560), 418-448.

Banks, J., & Oldfield, Z. (2007). Understanding pensions: Cognitive function, numerical ability and retirement saving. *Fiscal Studies*, *28*(2), 143-170.

Batty, G. D., Deary, I. J., & Zaninotto, P. (2016). Association of cognitive function with cause-specific mortality in middle and older age: follow-up of participants in the English Longitudinal Study of Ageing. *American Journal of Epidemiology*, *183*(3), 183-190.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4, Part 1), 352-365.

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences*, *109*(2), 425-430.

Brookmeyer, R., Gray, S., & Kawas, C. (1998). Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *American Journal of Public Health*, *88*(9), 1337-1342.

Caplan, Z. (2023, May 25). US older population grew from 2010 to 2020 at fastest rate since 1880 to 1890. *U.S. Census Bureau*. Retrieved from https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html

Card, D. (1999), "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card *(Eds.) Handbook of Labor Economics* (Vol. 3A), Amsterdam: North Holland, pp. 1801–1863.

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*(Suppl. 1), 123-160.

Cave, S. N., Wright, M., & von Stumm, S. (2022). Change and stability in the association of parents' education with children's intelligence. *Intelligence*, *90*, 101597. https://doi.org/10.1016/j.intell.2021.101597.

Chernozhukov, V., Lee, S., & Rosen, A. M. (2013). Intersection bounds: estimation and inference. *Econometrica*, *81*(2), 667-737.

Clark, D., and H. Royer (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review*, 103 (6): 2087-2120.

Courtin, E., Nafilyan, V., Glymour, M., Goldberg, M., Berr, C., Berkman, L. F., ... & Avendano, M. (2019). Long-term effects of compulsory schooling on physical, mental and cognitive ageing: a natural experiment. *Journal of Epidemiology & Community Health*, *73*(4), 370-376.

de Haan, M. (2011). The effect of parents' schooling on child's schooling: a nonparametric bounds analysis. *Journal of Labor Economics*, *29*(4), 859-892

Dickson, M., Gregg, P., & Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes?. *Economic Journal*, *126*(596), F184-F231.

Donovan, N. J., Wu, Q., Rentz, D. M., Sperling, R. A., Marshall, G. A., & Glymour, M. M. (2017). Loneliness, depression and cognitive function in older US adults. *International Journal of Geriatric Psychiatry*, *32*(5), 564-573.

Fletcher, J., Topping, M., Zheng, F., & Lu, Q. (2021). The effects of education on cognition in older age: Evidence from genotyped Siblings. *Social Science & Medicine*, *280*, 114044. https://doi.org/10.1016/j.socscimed.2021.114044.

Garrett, S. L., Kennedy, R. E., Sawyer, P., Williams, C. P., Brown, C. J., & Allman, R. M. (2019). Association between Executive Dysfunction and Instrumental Activities of Daily Living: Racial and ethnic differences among community-dwelling older adults in the Southeastern US. *Journal of the National Medical Association*, *111*(3), 320-327.

Germinario, G., Flores, C.A., & Flores-Lagunes, A. (2021). mpclr: A Stata command for bias-corrected estimation and valid inference for nonparametric bounds on treatment effects. Unpublished manuscript: Department of Economics, Syracuse University.

Glymour, M. M., Kawachi, I., Jencks, C. S., & Berkman, L. F. (2008). Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. *Journal of Epidemiology & Community Health*, *62*(6), 532-537.

Gorman, E. (2023). Does schooling have lasting effects on cognitive function? Evidence from compulsory schooling laws. *Demography*, *60*(4), 1139-1161.

Greenfield, E. A., & Moorman, S. M. (2019). Childhood socioeconomic status and later life cognition: Evidence from the Wisconsin Longitudinal Study. *Journal of Aging and Health*, *31*(9), 1589-1615.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy,* 80(2), 223–255.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65-73.

Herd, P., & Sicinski, K. (2022). Using sibling models to unpack the relationship between education and cognitive functioning in later life. *Social Science & Medicine-Population Health*, *17*, 100960. https://doi.org/10.1016/j.ssmph.2021.100960.

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, *49*(3), 615-651.

Hudomiet, P., Hurd, M. D., & Rohwedder, S. (2022). Trends in inequalities in the prevalence of dementia in the United States. *Proceedings of the National Academy of Sciences*, *119*(46), e2212205119. https://doi.org/10.1073/pnas.2212205119.

Kamhöfer, D. A., Schmitz, H., & Westphal, M. (2019). Heterogeneity in marginal non-monetary returns to higher education. *Journal of the European Economic Association*, *17*(1), 205-244.

Lachman, M. E., Agrigoroaei, S., Tun, P. A., & Weaver, S. L. (2014). Monitoring cognitive functioning: Psychometric properties of the brief test of adult cognition by telephone. *Assessment*, *21*(4), 404-417.

Lager, A., Seblova, D., Falkstedt, D., & Lövdén, M. (2017). Cognitive and emotional outcomes after prolonged education: a quasi-experiment on 320 182 Swedish boys. *International Journal of Epidemiology*, *46*(1), 303-311.

Lee, Y., Kim, J. H., Lee, K.J., Han, G., & Kim, J. L. (2005). Association of cognitive status with functional limitation and disability in older adults. *Aging Clinical and Experimental Research*, *17*, 20-28.

Liu, H., Chopik, W. J., Shrout, M. R., & Wang, J. (2024). A national longitudinal dyadic analysis of spousal education and cognitive decline in the United States. *Social Science & Medicine*, *343*, 116603. https://doi.org/10.1016/j.socscimed.2024.116603.

Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., ... & Mukadam, N. (2024). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*, *404*(10452), 572-628.

Magnuson, K. (2007). Maternal education and children's academic achievement during middle childhood. *Developmental Psychology*, *43*(6), 1497-1512.

Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables with an application to the returns to schooling. *Econometrica,* 68, 997-1010

Montez, J. K., Hummer, R. A., & Hayward, M. D. (2012). Educational attainment and adult mortality in the United States: A systematic analysis of functional form. *Demography*, *49*(1), 315-336.

Mullahy, J., Venkataramani, A., Millimet, D.L. & Manski, C.F. (2021). Embracing uncertainty: The value of partial identification in public health and clinical research. *American Journal of Preventive Medicine*, 61, e103–e108. https://doi.org/10.1016/j.amepre.2021.01.041.

Nafilyan, V., Avendano Pabon, M., & de Coulon, A. (2021). The Causal Impact of Depression on Cognitive Functioning: Evidence from Europe. IZA Discussion Paper No. 14049. Retrieved from https://docs.iza.org/dp14049.pdf.

Nicholas, L. H., Langa, K. M., Bynum, J. P.W., & Hsu, J. W. (2021). Financial presentation of Alzheimer disease and related dementias. *JAMA Internal Medicine*, *181*(2), 220-227.

Opdebeeck, C., Martyr, A., & Clare, L. (2016). Cognitive reserve and cognitive function in healthy older people: a meta-analysis. *Aging, Neuropsychology, and Cognition*, *23*(1), 40-60.

Sacerdote, B. (2002). The nature and nurture of economic outcomes. *American Economic Review: Papers and Proceedings*, *92*(2), 344-348.

Saenz, J. L., Beam, C. R., & Zelinski, E. M. (2020). The association between spousal education and cognitive ability among older Mexican adults. *The Journals of Gerontology: Series B*, *75*(7), e129-e140. https://doi.org/10.1093/geronb/gbaa002.

Schneeweis, N., Skirbekk, V., & Winter-Ebmer, R. (2014). Does education improve cognitive performance four decades after school completion? *Demography*, *51*(2), 619-643.

Stephens M. Jr, & Yang, D.Y. (2014). Compulsory education and the benefits of schooling. *American Economic Review*, *104*(6), 1777-1792.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(3), 448-460.

Stern, Y., Albert, M., Barnes, C. A., Cabeza, R., Pascual-Leone, A., & Rapp, P. R. (2023). A framework for concepts of reserve and resilience in aging. *Neurobiology of aging*, *124*, 100-103.

Xiao, Y., Li, L., & Zhao, L. (2017). Education on the cheap: The long-run effects of a free compulsory education reform in rural China. *Journal of Comparative Economics*, *45*(3), 544-562.

Xu, M. (2020). Spousal education and cognitive functioning in later life. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *75*(7), e141-e150. https://doi.org/10.1093/geronb/gbz014

Zahodne, L. B., Mayeda, E. R., Hohman, T. J., Fletcher, E., Racine, A. M., Gavett, B., ... & Mungas, D. (2019). The role of education in a vascular pathway to episodic memory: brain maintenance or cognitive reserve? *Neurobiology of Aging*, *84*, 109-118.

Figure 1: Conceptual Framework Linking Schooling to Cognition at Older Ages

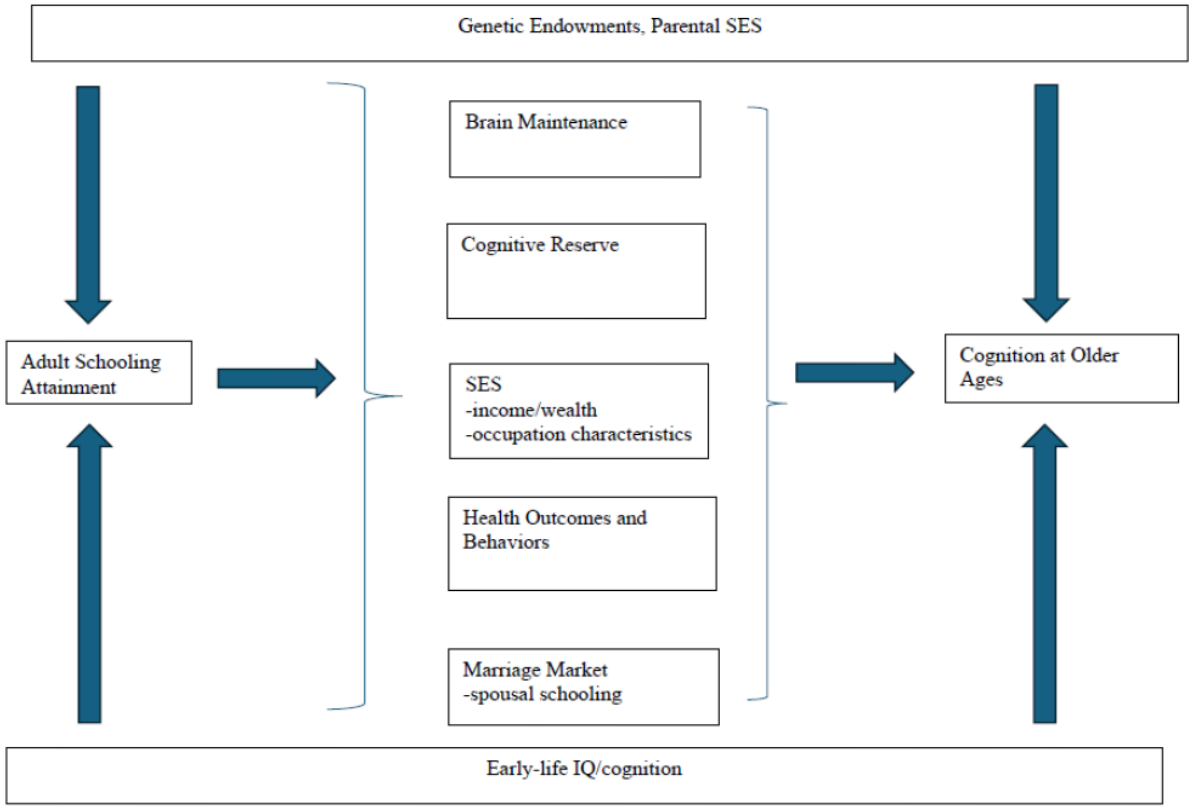
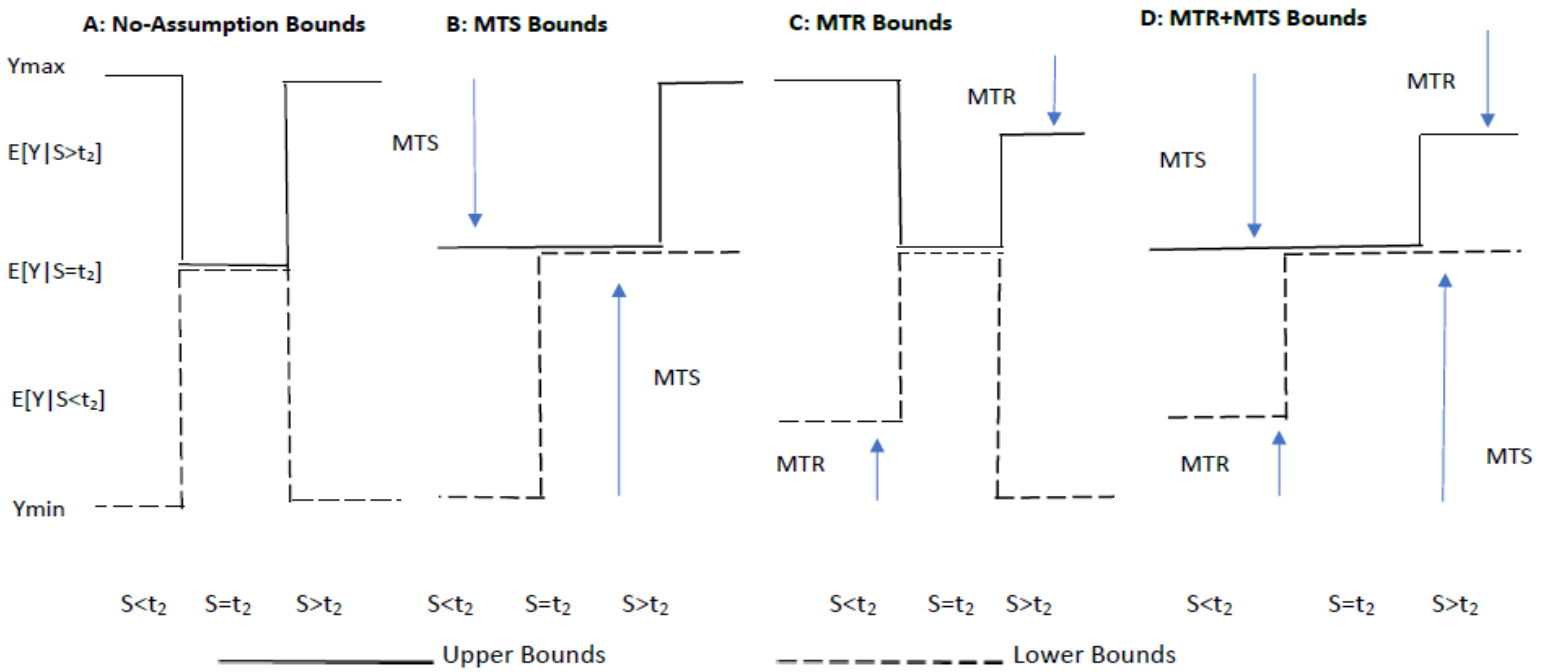


Figure 2: No-Assumption, MTS and MTR Bounds for $E[Y(t_2)]$



Notes: Each panel illustrates graphically the construction of nonparametric bounds on E[Y(t_2)] under a given assumption. The x-axis represents levels of the realized schooling (S) while the y-axis represents values of the outcome (Y). Panel A illustrates the construction of the no-assumption bounds where the maximum (Ymax) and minimum (Ymin) of the outcome are used to replace the missing counterfactuals for those with S<t_2 and S >t_2. In panel B, the MTS assumption tightens the lower and upper bounds relative to the no-assumption bounds by using the observed mean outcomes for those receiving t_2 (E[Y| S=t_2]) to replace the missing counterfactuals for those with S<t_2 and S>t_2. In panel C, the MTR assumption reduces the upper bound for those with S>t_2 by using the observed mean outcome for those with S>t_2 (E[Y|S>t_2]). The MTR assumption increases the lower bound for those with S<t_2 by using the observed mean outcome for those with S<t_2 (E[Y|S<t2]). The combination of panel B and panel C would correspond to the case when the MTS and MTR assumptions are imposed together in panel D.

Table 1: Summary Statistics

	Full Sample	Men	Women	Whites	Blacks	Hispanics
Demographics						
Age	76.15 (7.41)	75.96 (7.18)	76.27 (7.55)	76.80 (7.47)	74.41 (6.97)	74.36 (6.98)
Female	0.61 (0.49)	—	—	0.59 (0.49)	0.69 (0.46)	0.61 (0.49)
Race: Non-Hispanic White	0.73 (0.44)	0.77 (0.42)	0.71 (0.45)	—	—	—
Race: Non-Hispanic Black	0.16 (0.35)	0.12 (0.32)	0.16 (0.37)	—	—	—
Race: Hispanic	0.10 (0.30)	0.10 (0.30)	0.10 (0.32)	—	—	—
Race: Other	0.02 (0.15)	0.02 (0.13)	0.02 (0.15)			
Mother: HS Dropout	0.54 (0.50)	0.48 (0.50)	0.58 (0.49)	0.46 (0.50)	0.68 (0.47)	0.87 (0.33)
Mother: HS Grad	0.32 (0.47)	0.37 (0.48)	0.28 (0.45)	0.36 (0.48)	0.25 (0.44)	0.09 (0.28)
Mother: More than HS	0.14 (0.35)	0.14 (0.35)	0.14 (0.35)	0.17 (0.38)	0.06 (0.24)	0.04 (0.20)
Schooling						
Grades of Schooling	12.92 (3.07)	13.24 (3.06)	12.71 (3.05)	13.49 (2.44)	12.50 (2.84)	9.35 (4.55)
HS Dropout	0.18 (0.38)	0.16 (0.36)	0.19 (0.40)	0.11 (0.32)	0.24 (0.43)	0.55 (0.50)
HS Grad	0.33 (0.47)	0.31 (0.46)	0.34 (0.47)	0.35 (0.48)	0.33 (0.47)	0.20 (0.40)
Some College	0.23 (0.42)	0.22 (0.41)	0.24 (0.43)	0.24 (0.43)	0.23 (0.42)	0.16 (0.37)
College Grad	0.25 (0.44)	0.31 (0.46)	0.22 (0.42)	0.29 (0.46)	0.19 (0.40)	0.08 (0.27)
Cognition						
Mini-Mental State Exam	26.76 (3.84)	26.60 (3.64)	26.87 (3.95)	27.26 (3.48)	25.48 (4.22)	25.27 (4.61)
Immediate and Delayed Memory	48.81 (18.12)	46.46 (16.90)	50.33 (18.72)	50.72 (17.98)	44.33 (17.95)	42.06 (16.49)
Memory Recognition	28.82 (4.44)	28.56 (4.15)	28.99 (4.60)	29.25 (4.26)	28.13 (4.89)	26.88 (4.34)
Verbal Fluency	16.16 (6.65)	16.45 (6.68)	15.96 (6.62)	16.92 (6.69)	13.55 (6.19)	14.71 (5.76)
Attention/Speed	44.02 (15.08)	44.97 (14.85)	43.40 (15.20)	43.36 (13.99)	38.05 (15.67)	36.14 (16.15)
Executive Function	12.44 (3.73)	12.84 (3.73)	12.17 (3.70)	13.19 (3.33)	9.98 (4.00)	10.70 (3.83)
Visuospatial ability	8.12 (2.33)	8.42 (2.25)	8.07 (2.37)	8.52 (2.22)	7.15 (2.36)	7.41 (2.51)
Observations	3,072	1,206	1,866	2,248	448	310

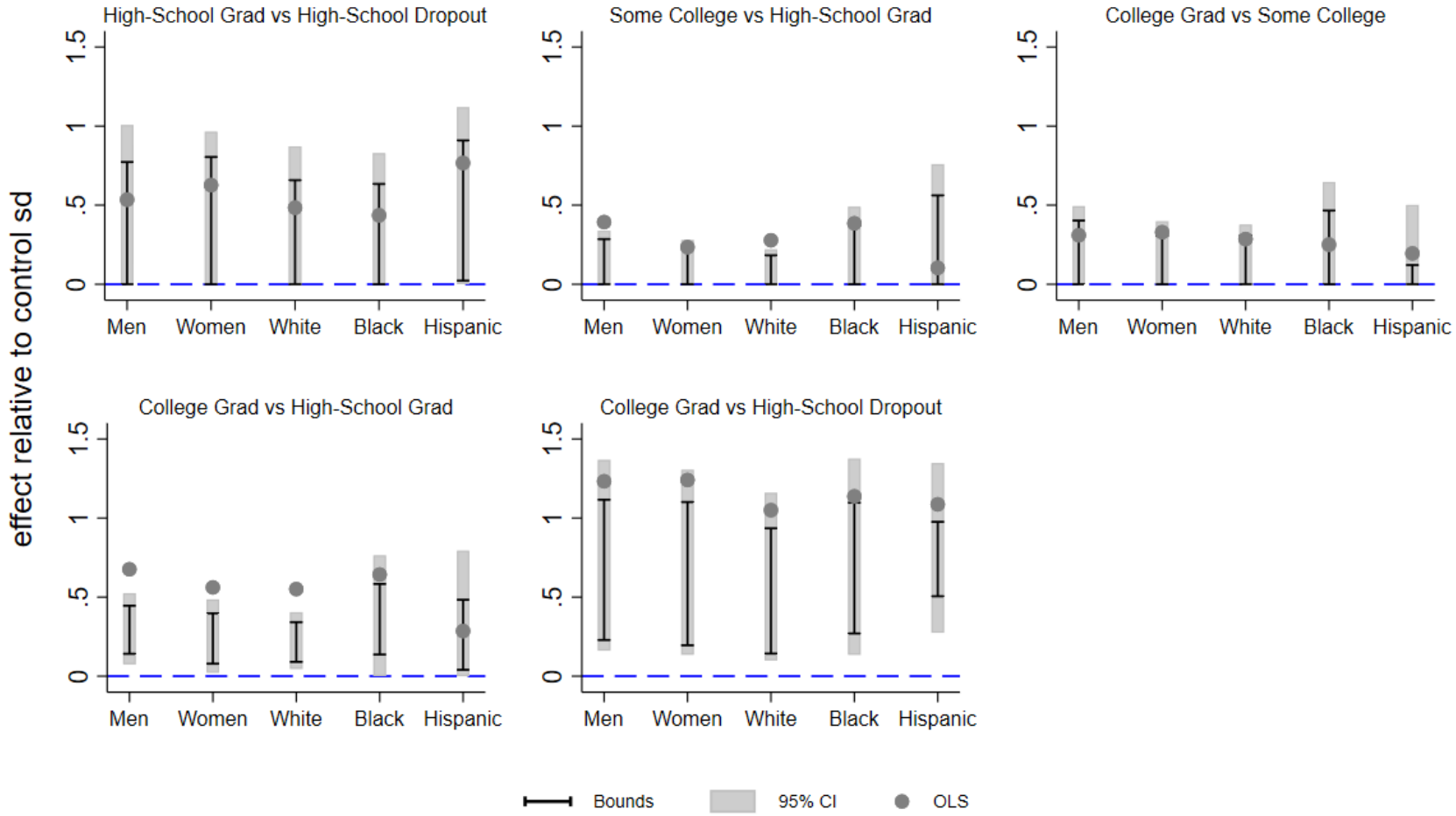
Notes: Standard deviations are shown in parentheses.

Table 2: OLS Estimates and Bounds for the Effect of Schooling on Immediate and Delayed Memory

	Control Mean (SD)	OLS	No Assumption	MTS	MTR	MTR+MTS	MTS+MTR+MIV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A							
HS Grad vs HS Dropout	37.55 (15.95)	9.29*** (0.87)	[-71.80, 74.41] (-72.61, 75.21)	[-48.75,34.47] (-49.52,35.90)	[0.00, 52.89] (0.00, 53.40)	[0.00, 12.92] (0.00, 14.18)	[0.00, 11.48] (0.00, 13.42)
Panel B							
Some College vs HS Grad	46.82 (17.22)	4.90*** (0.82)	[-69.07, 71.78] (-69.94 ,72.55)	[-46.55, 25.24] (-47.23, 26.47)	[0.00, 54.33] (0.00, 54.85)	[0.00, 7.79] (0.00, 8.94)	[0.00, 4.10] (0.00, 4.71)
Panel C							
College vs Some College	51.73 (16.55)	4.73*** (0.88)	[-72.59, 75.18] (-73.41, 75.93)	[-49.09, 31.01] (-49.88, 32.32)	[0.00, 53.06] (0.00, 53.57)	[0.00, 8.88] (0.00, 10.09)	[0.00, 5.26] (0.00, 6.32)
Panel D							
College vs HS Grad	46.82 (17.22)	9.64*** (0.83)	[-66.50, 71.80] (-67.35, 72.56)	[-57.39, 17.98] (-58.22, 19.37)	[0.00, 65.11] (0.00, 65.84)	[0.00, 11.29] (0.00, 12.63)	[1.87, 6.74] (1.14, 7.75)
Panel E							
College vs HS Dropout	37.55 (15.95)	18.91*** (0.93)	[-72.61, 80.55] (-73.54, 81.24)	[-72.61, 18.92] (-73.54, 20.42)	[0.00, 80.55] (0.00, 81.24)	[0.00, 18.92] (0.00, 20.44)	[3.43, 16.30] (2.75, 18.46)

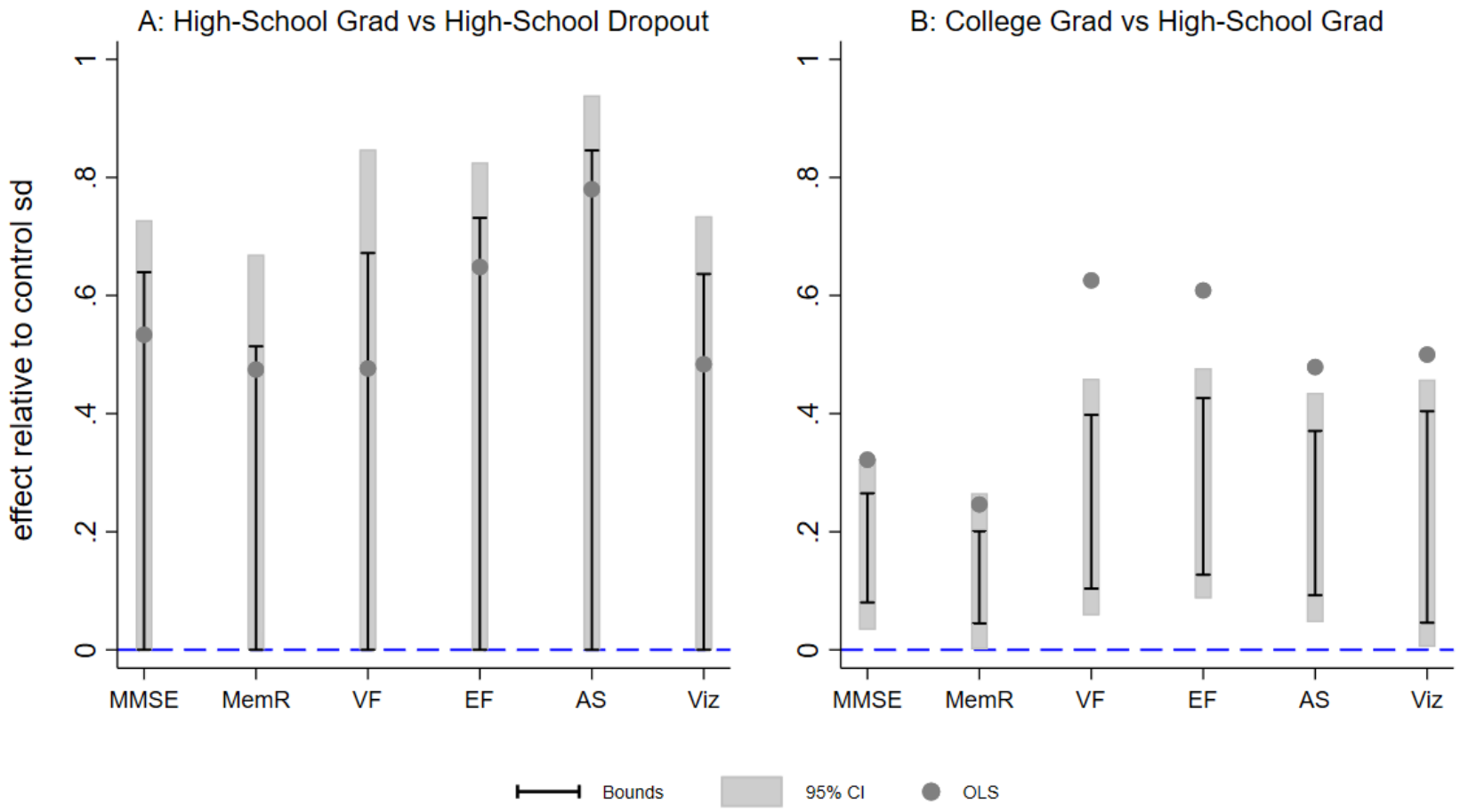
Notes: Robust standard errors in (.) in column 2. In columns 3-7 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome was used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school.
***$p<0.001$

Figure 3: MTS+MTR+MIV Bounds on Effects of Schooling on Immediate+Delayed Memory By Gender and Race



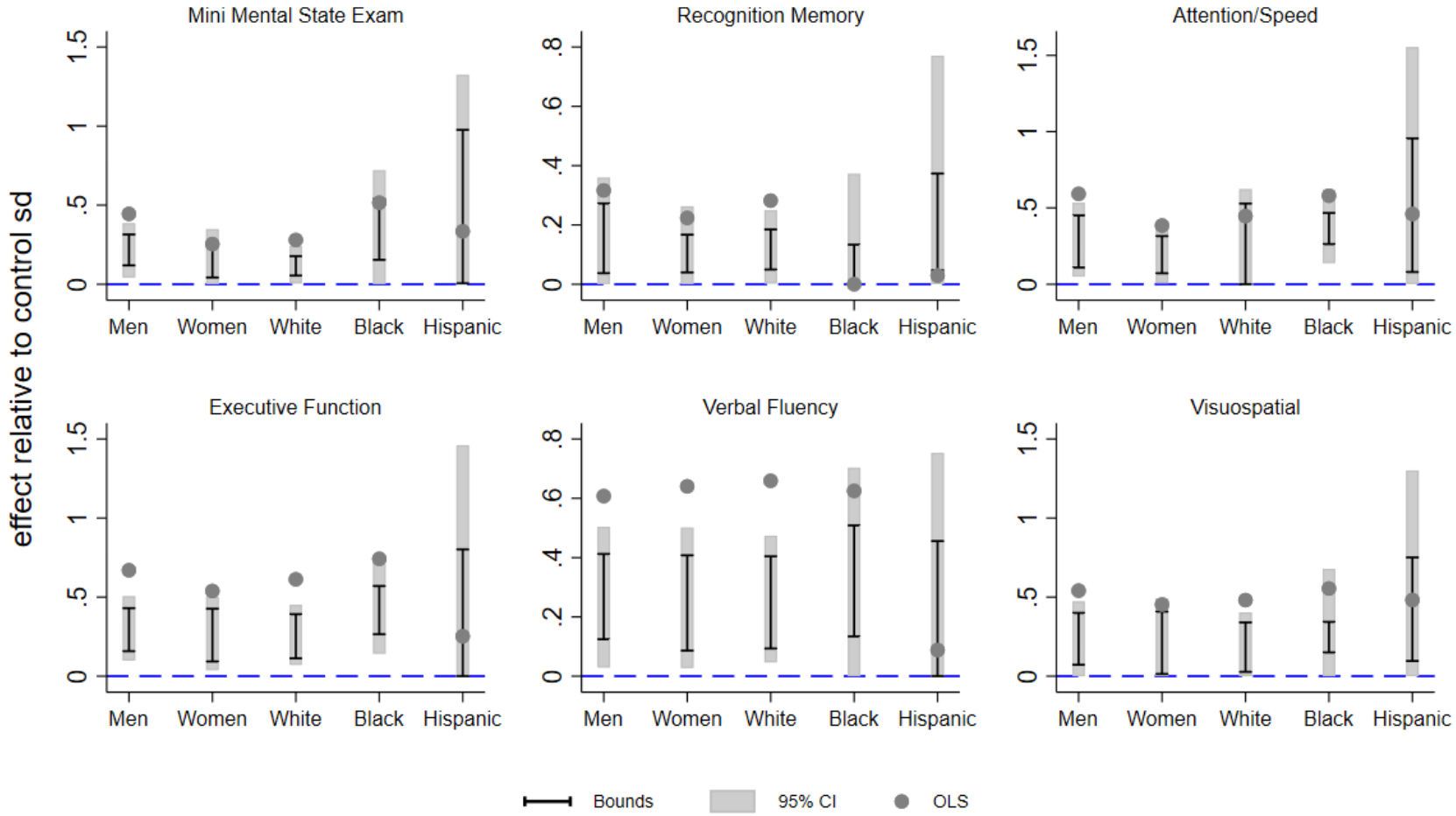
Notes: The figure presents the estimated bounds on the effects of increasing schooling at different margins on immediate and delayed memory, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A.

Figure 4: MTS+MTR+MIV Bounds For Other Cognition Domains



Notes: The figure presents the estimated bounds on the effects of completing secondary schooling in panel A and increasing schooling from secondary to tertiary schoolin in panel B. The y-axis indicates the size of the effects relative to the corresponding "control" sd. The control group refers to high-school dropouts in panel A and high-school graduates in panel B. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A. MMSE:Mini Mental State Examination; MemR:Recognition Memory; VF:Verbal Fluency; EF:Executive Function; AS:Attention/Speed Viz:Visuospatial

Figure 5: MTS+MTR+MIV Bounds on Effect of Increasing Schooling from High School to College on Other Cognition Domains By Gender and Race



Notes: The figure presents the estimated bounds on the effect of going from being a high-school graduate to college graduate on cognition, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control (high school graduates)" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described appendix A.

Table 3: Bounds for the Effects of Schooling on Health, Poverty, Occupation, and Spousal Schooling at Older Ages

Panel A: Health Measures	**Poor or Fair Health**	**BMI**	**CES-D**	**Ever-Smoke**	**High blood pressure**	**Not engaging in vigorous exercise**
HS Grad vs HS Dropout	[-0.203, 0.000] (-0.254, 0.000)	[-0.196, 0.000] (-0.956, 0.000)	[-0.830, 0.000] (-1.031, 0.000)	[-0.048, -0.007] (-0.093, 0.000)	[-0.116, 0.000] (-0.149, 0.000)	[-0.103, 0.000] (-0.151, 0.000)
Some College vs HS Grad	[-0.077, 0.000] (-0.095, 0.000)	[-0.164, -0.007] (-0.460, 0.000)	[-0.200, 0.000] (-0.294, 0.000)	[-0.012, -0.013] (-0.048, 0.000)	[-0.042, 0.000] (-0.061, 0.000)	[-0.031, 0.000] (-0.053, 0.000)
College Grad vs Some College	[-0.078, 0.000] (-0.105, 0.000)	[-0.370, 0.000] (-0.894, 0.000)	[-0.229, 0.000] (-0.362, 0.000)	[-0.046, -0.002] (-0.087, 0.000)	[-0.065, 0.000] (-0.093, 0.000)	[-0.135, -0.008] (-0.167, 0.000)
College Grad vs HS Grad	[-0.114, -0.021] (-0.139, -0.000)	[-0.443, 0.000] (-0.928, 0.000)	[-0.343, -0.012] (-0.480, 0.000)	[-0.059, -0.021] (-0.097, 0.000)	[-0.075, -0.011] (-0.102, 0.000)	[-0.141, -0.027] (-0.171, 0.000)
College Grad vs HS Dropout	[-0.270, -0.058] (-0.324, -0.039)	[-0.553, 0.000] (-1.374, 0.000)	[-1.014, -0.135] (-1.251, -0.045)	[-0.109, -0.027] (-0.161, 0.000)	[-0.172, -0.022] (-0.213, -0.003)	[-0.234, -0.032] (-0.291, -0.005)
Panel B: SES Measures	**In Poverty**	**Professional/ Managerial Occupations**	**Spouse is a College Graduate**			
HS Grad vs HS Dropout	[-0.124, 0.000] (-0.158, 0.000)	[0.000, 0.122] (0.000, 0.152)	[0.000, 0.288] (0.000, 0.323)			
Some College vs HS Grad	[-0.044, 0.000] (-0.056, 0.000)	[0.000, 0.040] (0.000, 0.062)	[0.000, 0.096] (0.000, 0.115)			
College Grad vs Some College	[-0.037, 0.000] (-0.052, 0.000)	[0.002, 0.080] (0.000, 0.107)	[0.000, 0.272] (0.000, 0.304)			
College Grad vs HS Grad	[-0.057, -0.013] (-0.072, -0.001)	[0.022, 0.090] (0.003, 0.119)	[0.062, 0.284] (0.041, 0.316)			
College Grad vs HS Dropout	[-0.155, -0.034] (-0.192, -0.021)	[0.031, 0.194] (0.013, 0.235)	[0.076, 0.548] (0.056, 0.587)			

Notes: Estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome were used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school. For the variable "Spouse is a College Graduate", the analysis is done only for HCAP respondents who are part of a coupled household. We use information on spousal schooling from the HRS 2016 wave.

Does Schooling Improve Cognitive Abilities at Older Ages? Causal Evidence from Nonparametric Bounds

Online Appendix

Appendix A: Technical Appendix

This appendix provides additional details on our econometric approach. Let every individual i have a response function $Y(\cdot): T \to \mathcal{Y}$ which maps treatments $t \in T$ into potential outcomes $Y_i(t) \in \mathcal{Y}$. In our context, the treatment t is schooling attainment consisting of four levels: high-school dropouts (t_1), high-school graduates (t_2), some college education (t_3) and college graduates (t_4). Let S_i denote the realized treatment received by individual i, so that $Y_i \equiv \sum_{t \in T} 1\{S_i = t\} \cdot Y_i(t)$ is the associated observed outcome, where $1\{A\}$ is the indicator function which equals one if the statement A is true and equals zero otherwise. We are interested in the population ATE of, for example, increasing schooling attainment from t_1 to t_2 on cognition test scores, defined as:

$$\text{(A1)} \quad \Delta(t_1, t_2) = E[Y(t_2)] - E[Y(t_1)]$$

As discussed in section 3 of the paper, estimation of the ATE is complicated because of the missing counterfactuals, which can be seen by using the law of iterated expectations to write the expected potential outcome $E[Y(t_2)]$ as:

$$\begin{aligned}\text{(A2)}\ E[Y(t_2)] = {} & E[Y(t_2)|S < t_2] * P(S < t_2) + E[Y(t_2)|S = t_2] * P(S = t_2) \\ & + E[Y(t_2)|S > t_2] * P(S > t_2)\end{aligned}$$

Manski (1989) suggested a bounded-support assumption, whereby one uses the minimum (Y_{min}) and maximum (Y_{max}) of the outcome variable in place of the counterfactuals. This gives Manski's (1989) "no-assumption" bounds:

$$\text{(A3) } Y_{min} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S > t_2) \leq E[Y(t_2)] \leq Y_{max} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S > t_2)$$

The no-assumption lower (upper) bound on the ATE $\Delta\,(t_1, t_2)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$. Bounds for other treatment effects such as $\Delta\,(t_2, t_3)$, $\Delta\,(t_1, t_3)$ or $\Delta\,(t_2, t_4)$ are computed analogously. In practice, the no-assumption bounds are typically wide and uninformative, and contain zero by construction. To tighten the bounds, we employ three monotonicity assumptions introduced in Manski (1997) and Manski and Pepper (2000): (1) monotone treatment selection; (2) monotone treatment response; and (3) monotone instrumental variable.

Monotone Treatment Selection (MTS)

We employ the non-negative MTS assumption which captures the notion that, on average, individuals who "selected" into higher education have higher latent cognitive abilities. Formally, for each $t \in T$ and two treatment levels μ_1 and μ_2

$$\text{(A4) } \mu_2 \geq \mu_1 \Rightarrow E[Y(t) \mid S = \mu_2] \geq E[Y(t) \mid S = \mu_1]$$

The MTS assumption requires that individuals with higher schooling attainment on average have weakly higher potential outcomes at every schooling level t. For example, when comparing high-

school graduates (e.g., $S = \mu_2$) to high-school dropouts ($S = \mu_1$), the MTS assumption requires that the average potential cognition at older ages at any schooling level t (e.g., as a college graduate) of high-school graduates is higher than that of high-school dropouts.

Consider the conditional mean potential outcomes for individuals with $S < t_2$ and $S > t_2$. Under the MTS assumption, $E[\,Y(t_2) \mid S < t_2\,]$ cannot be more than $E[\,Y(t_2) \mid S = t_2\,]$, which is identified by the observed mean outcome for those receiving t_2. The observed mean outcome for those receiving t_2 can therefore be used as an upper bound for the mean of $Y(t_2)$ for those with $S < t_2$. Similarly, for the conditional mean potential outcome $E[\,Y(t_2) \mid S > t_2\,]$, the MTS implies that the unidentified quantity cannot be smaller than $E[\,Y(t_2) \mid S = t_2\,]$, or the observed mean outcome for $S = t_2$, $E[Y|S = t_2]$. This implies that the observed mean outcome for those receiving t_2 can be used as a lower bound for the mean of $Y(t_2)$ for those with $S > t_2$. Then, the MTS bounds on $E[Y(t_2)]$ are given by (Manski and Pepper, 2000):

$$\begin{aligned}(\text{A5})\ Y_{min} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2) \\ \leq E[Y(t_2)] \leq \\ E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S > t_2)\end{aligned}$$

As before, the lower (upper) bound on the ATE $\Delta\,(t_1, t_2)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$, and likewise for other comparisons of interest.

Monotone Treatment Response (MTR)

We employ the non-negative MTR assumption, imposing the restriction that higher schooling attainment does not decrease cognitive ability at older ages. Formally, for each individual and any treatment levels t_k and t_j:

(A6) $t_j \geq t_k \Longrightarrow Y(t_j) \geq Y(t_k)$

A key implication from MTR is that, for example, $E[Y(t_2)|S = t_\ell] \geq E[Y(t_1)|S = t_\ell]$ for any ℓ, given that $t_2 > t_1$. For any treatment levels $t < t_2$, MTR implies that the conditional mean $E[Y(t_2)|S = t]$ is no less than $E[Y(t)|S = t]$, or the observed mean of Y at t, $E[Y|S = t]$. This increases the lower bound on $E[Y(t_2)]$, relative to that obtained from the bounded support assumption alone. Further, for treatment levels $t' > t_2$, MTR implies that the conditional mean $E[Y(t_2)|S = t']$ cannot be more than $E[Y(t')|S = t']$, which is identified by the observed mean of Y at t', $E[Y|S = t']$. This reduces the upper bound on the unconditional mean $E[Y(t_2)]$ when compared to the no-assumption upper bound.

The MTR bounds on $E[Y(t_2)]$ are given by (Manski 1997):

$$\begin{aligned}(\text{A7})\ & E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S > t_2) \\ & \leq E[Y(t_2)] \leq \\ & Y_{max} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2)\end{aligned}$$

As usual, the MTR lower (respectively, upper) bound on the ATE $\Delta(t_1, t_2)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$. Under the non-negative MTR assumption the lower bound on $\Delta(t_1, t_2)$ is never below zero, because the MTR rules out the possibility that more education worsens cognitive abilities.

Montone Treatment Selection Combined with Monotone Treatment Response

The MTR and MTS assumptions can be combined to provide tighter bounds on mean potential outcomes. The MTS+MTR bounds on $E[Y(t_2)]$ are given by (Manski and Pepper, 2000):

(A8) $E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2) \leq E[Y(t_2)] \leq E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2)$

As before, the MTR+MTS lower (upper) bound on the ATE $\Delta(t_1, t_2)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$. The MTS and MTR assumptions imposed together yield a testable implication that observed mean cognition scores are weakly increasing in schooling attainment. That is, for any two treatments t_k and t_j, $t_j > t_k$ implies that $E[Y|S = t_j] \geq E[Y|S = t_k]$. This is the case because $t_j > t_k$ implies:

(A9) $E[Y|S = t_j] = E[Y(t_j)|S = t_j] \geq E[Y(t_j)|S = t_k] \geq E[Y(t_k)|S = t_k] = E[Y|S = t_k]$,

where the first inequality follows from the MTS assumption and the second from the MTR. Lastly, note that these inequalities help to highlight a key distinction between the MTS and MTR assumptions: while the MTS compares the mean of the same potential outcome for two different subpopulations defined by their observed levels of t, the MTR compares different potential outcomes for the same individual(s). For example, the MTS would compare the average potential cognition at a given schooling level (e.g., as a college graduate) of high-school graduates versus high-school dropouts, while the MTR would compare the potential cognition under high-school graduation versus the potential cognition under a high-school dropout for the same individual(s).

Monotone Instrumental Variable (MIV)

A MIV is a variable that has a monotone (weakly increasing or weakly decreasing) mean relationship with the potential outcomes $Y(t)$. Specifically, a weakly increasing MIV Z satisfies:

(A10) $m_1 \leq m \leq m_2 \Rightarrow E[Y(t)|Z = m_1] \leq E[Y(t)|Z = m] \leq E[Y(t)|Z = m_2]$

for all treatment levels $t \in T$.

To motivate how the MIV narrows bounds, consider using it along with the MTS+MTR bounds. With a variable Z satisfying the MIV assumption, we can divide the sample into bins defined by the values of Z and compute the MTS+MTS bounds within each bin. We use mother's schooling as the MIV. In our case of a non-negative MIV, equation (A10) implies that the lower bound on $E[Y(t_2)|Z = m]$ is no lower than the lower bound on $E[Y(t_2)|Z = m_1]$, and its upper bound is no higher than the upper bound on $E[Y(t_2)|Z = m_2]$. For the bin where Z has a value of m, we can thus obtain a new lower bound by taking the largest lower bound over all bins where $Z \leq m$. This type of bounding approach is known as intersection-bounds. This is illustrated in Figure A1, which shows fictitious MTS+MTR upper and lower bounds in three bins of Z. In the subsample Z=2, the maximum lower bound is estimated over all subsamples with Z≤ 2. This is the lower bound at Z=1, so this becomes the new lower bound at Z=2. Likewise, we can obtain a new upper bound by taking the smallest upper bound over all bins where $Z \geq m$. Looking at Z=2 in Figure A1, the lowest upper bound over all values Z≥ 2 turns out to be the upper bound at Z=3. So, this becomes the new upper bound at Z=2. The MTS+MTR+MIV bounds are then obtained by taking the weighted average over all the conditional-on-Z bounds (which follows from the law of iterated expectations):

(A11) $\sum_{m \in M} P(Z = m) * \left[max_{m_1 \leq m} LB_{E[Y(t_2)|Z=m_1]}\right]$

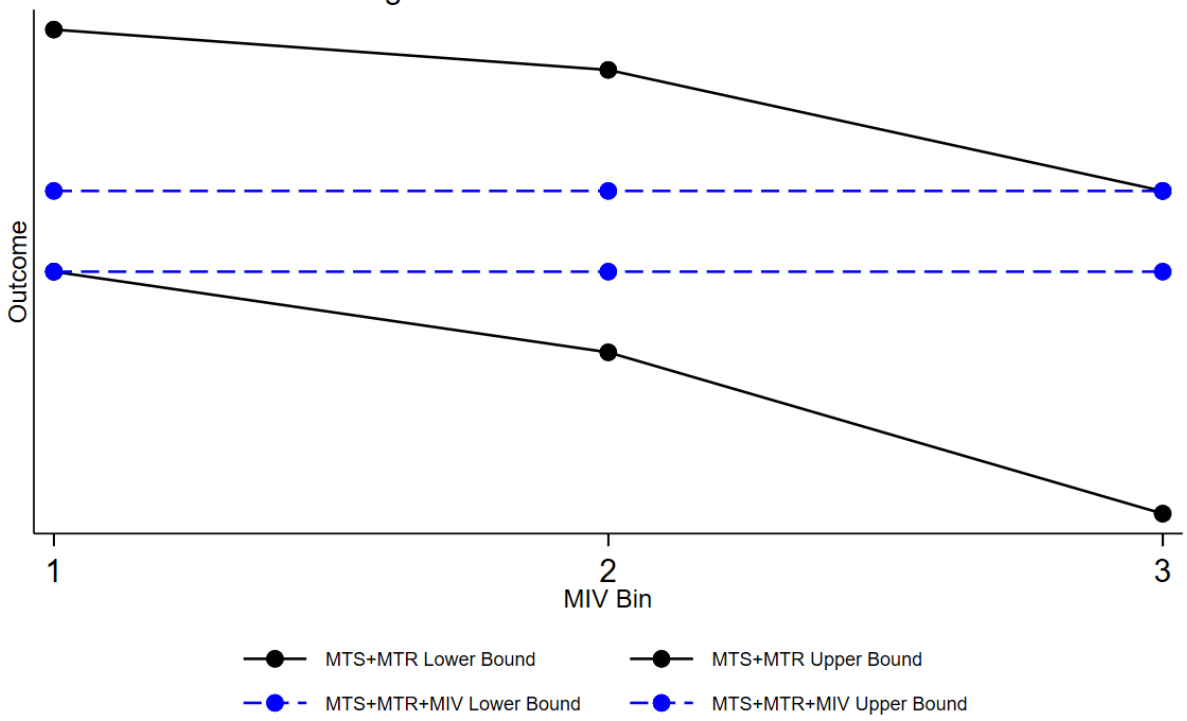
$\leq E[Y(t_2)] \leq$

$\sum_{m \in M} P(Z = m) * \left[min_{m_2 \geq m} UB_{E[Y(t_2)|Z=m_2]}\right]$

where LB denotes the MTS+MTR lower bound from equation (A8) on $E[Y(t_2)]$ at values $Z = m_1$ of the MIV. Likewise, UB represents the MTR+MTS upper bound on $E[Y(t_2)]$ conditional on values $Z = m_2$ of the MIV.

The MTR+MTS+MIV lower (upper) bound on the ATE $\Delta\,(t_1, t_2)$ is calculated once again by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$.

Figure A1: MTS+MTR+MIV Bounds



Notes: Construction of MTS+MTR+MIV bounds obtained with a weakly increasing MIV (Z) divided into 3 bins. For a bin where Z=m MTS+MTR+MIV lower bound is obtained by taking the largest MTS+MTR lower bound over all bins where Z≤m. MTS+MTR+MIV upper bound is obtained by taking the smallest MTS+MTR upper bound over all bins where Z≥m. Figure based on De Hann (2011).

Estimation and Inference Issues

As noted in the paper, all bounds are estimated by plugging in sample analogs for the expectations and probabilities in the corresponding bounds' expressions. Inference is undertaken

by constructing confidence intervals that cover the true value of the population average treatment effect of interest with a specified probability (e.g., 95%).

Estimation and inference under the MTR+MTS+MIV bounds require that we deal with two issues that have been noted since Manski and Pepper (2000)—see, e.g., Tamer (2010) and references therein. The first is that the plug-in estimators of equation (A11)—an example of so-called intersection bounds—suffer from bias in finite samples that makes them narrower relative to the corresponding true identified set. The bias then carries over to estimated bounds on the average treatment effects of interest. The second, related issue is that the corresponding confidence intervals do not have the expected coverage at the desired level. Both of these issues arise because of the non-concavity and non-convexity, respectively, of the min and max operators in equation (A11).

We address both issues in the bounds involving the MIV assumption by employing the estimation and valid-inference procedure in Chernozhukov et al. (2013; hereafter, CLR) for intersection bounds.[1] The CLR procedure allows us to obtain lower- and upper-bound estimators that satisfy a half-median unbiasedness property, that is, the estimated lower (upper) bound will fall below (above) the true lower (upper) bound with a probability of at least one-half asymptotically. This property is important because Hirano and Porter (2012) showed that there exist no locally asymptotically unbiased estimators of parameters that contain min and max operators, implying that methods aimed at reducing bias (such as those based on the bootstrap) cannot completely eliminate it and reducing bias too much eventually leads the variance of such methods to increase significantly.

[1] See also Flores and Flores-Lagunes (2013) for additional discussion on the CLR procedure and an application estimating bounds on local-average-treatment effects without the exclusion restriction under a different set of monotonicity assumptions.

Computation of all the estimated bounds and confidence intervals, including for cases not using the MIV assumption, is done by employing the CLR methodology. While the use of the CLR methodology is only strictly necessary for the MIV bounds (which contain maxima/minima), we use it for all bounds for the practical reason of not requiring a different procedure. Indeed, the CLR methodology can be thought of as a generalization of the methods in Imbens and Manski (2004) and Stoye (2009), which are commonly used to create confidence intervals on the parameter of interest for bounds that do not contain minimum or maximum operators (see, e.g., CLR, 2013, footnote 13). As for Stoye's (2009) confidence intervals, CLR's confidence intervals remain valid even when the width of the bounds is not bounded away from zero (CLR, 2009). Regarding the estimated bounds for non-intersection bounds, the use of the CLR procedure results in minuscule corrections, as expected, since they do not contain minimum or maximum operators (or, in the context of CLR's methodology, they take the minimum or maximum of a single object).

Implementation of the CLR Procedure

To provide some intuition on the CLR method, we first make explicit the notion of creating the bins of the MIV. We use below 3 MIV bins $\mathcal{B}_m$, $m = 1, 2, 3$. Then, for instance, the lower bound on $E[Y(t_2)]$ from equation (A11) can be rewritten as:

$$(\text{A12}) \sum_{m=1}^{3} P(Z \in \mathcal{B}_m) \cdot \max_{m_1 \leq m} LB^1_{m_1}$$

where the $LB^1_{m_1}$ are the MTR-MTS lower bounds in bins m_1 up through m.

Instead of expressions like (A12) which comprise 3 different maxima, the CLR method requires that these be rewritten as a set of expressions under a single maximum (or minimum, for upper bounds), with each element inside the max operator called a bounding function. These bounding functions are objects for which standard (e.g., asymptotically normally distributed) estimators exist, and hence they do not contain max/min operators. Intuitively, each bounding function represents one of the possible outcomes from evaluating (A12) in the data. To see that (A12) implies three different maxima, note that it can be written as: $P\left(Z \in \mathcal{B}_{m_1}\right) \cdot \max\{LB^1_{m_1}\} \quad + P\left(Z \in \mathcal{B}_{m_2}\right) \cdot \max\{LB^1_{m_1}, LB^2_{m_2}\} + P\left(Z \in \mathcal{B}_{m_3}\right) \cdot \max\{LB^1_{m_1}, LB^2_{m_2}, LB^3_{m_3}\}$. These expressions with different maxima can be manipulated using the properties of the maximum operator to obtain an expression with a single maximum operator over six bounding functions as follows. Let $p_l \equiv P(Z \in \mathcal{B}_l)$, then: $\max\{(p_1 + p_2 + p_3)\, LB^1_{m_1}, (p_1 + p_2)\, LB^1_{m_1} + p_3 LB^2_{m_2}, (p_1 + p_2)\, LB^1_{m_1} + p_3 LB^3_{m_3}, p_1 LB^1_{m_1} + p_2 LB^2_{m_2} + p_3 LB^1_{m_1}, p_1 LB^1_{m_1} + p_2 LB^2_{m_2} + p_3 LB^2_{m_2}, p_1 LB^1_{m_1} + p_2 LB^2_{m_2} + p_3 LB^3_{m_3}\}$.

Finally, the full set of bounding functions is defined for the ATE, so we also perform all necessary subtractions. For example, the final bounding functions for the lower bound on $\Delta\,(t_1, t_2)$ are created from all possible subtractions of the $E[y(t_1)]$ upper bound bounding functions from the $E[Y(t_2)]$ lower bound bounding functions. In total, each bound on each ATE implies (2^{3 − 1})^2 = 16 bounding functions, denoted $\theta^l(v)$ and $\theta^u(v)$, $v = 1, \ldots, 16$, for the respective lower and upper bounds. Using these bounding functions and denoting the true value of the lower bound of the ATE as θ_0^l and the one for the upper bound as θ_0^u, we can write

$$\theta_0^l = \max_{v \in \mathcal{V}^l} \{\theta^l(v)\}$$

and

$$\theta_0^u = \min_{v \in \mathcal{V}^u} \{\theta^u(v)\},$$

where $\mathcal{V}^l$ and $\mathcal{V}^u$ are the indexing sets for the bounding functions of the lower ($\theta^l(v)$) and upper ($\theta^u(v)$) bounds, respectively.

The key aspect of the CLR procedure is that the steps for estimation of the bounds and for constructing confidence intervals are completed on the individual bounding functions prior to taking the associated maximum (or minimum). This is referred to as the *precision adjustment* and proceeds as follows.[2] Generally, the adjustment involves taking the product of a critical value $\kappa(p)$ and the pointwise standard error $s(v)$ of the bounding function. For lower bounds, this product is subtracted from the estimator $\widehat{\theta^l}(v)$; for upper bounds, it is added to $\widehat{\theta^u}(v)$. Then—depending on the choice of critical value p—the adjustment yields either the half-median unbiased estimator of the lower and upper bounds ($p = 0.5$), or the desired lower and upper limits of the confidence interval (see below). In this way, the CLR method offers the convenience that median-bias correction and inference are carried out within the same procedure. Also, we note that the resulting large number of bounding functions makes it crucial to implement the CLR procedure for estimation of the bounds and the construction of valid confidence intervals, as in practice the amount of bias tends to increase with the number of bounding functions (e.g., Germinario et al. 2021).

More specifically, the precision-corrected estimators of the lower (θ_0^l) and upper (θ_0^u) bounds of the average treatment effect are given, respectively, by:

[2] This process requires that the estimators of $\theta^l(v)$ and $\theta^u(v)$ are consistent and asymptotically normal. Since in our case these estimators are made up of sample means and sample proportions, these conditions are met.

(A13) $\widehat{\theta^l}(p) = \max_{v \in \mathcal{V}^l} \{\widehat{\theta^l}(v) - \kappa^l(p) \cdot s^l(v)\}$

and

(A14) $\widehat{\theta^u}(\mathrm{p}) = \min_{v \in \mathcal{V}^u} \{\widehat{\theta^u}(\mathrm{v}) + \kappa^u(\mathrm{p}) \cdot \mathrm{s}^u(\mathrm{v})\}$

where $\widehat{\theta^l}(v)$ and $\widehat{\theta^u}(v)$ are the unadjusted estimators of the bounding functions, and $s^l(v)$ and $s^u(v)$ are their associated standard errors. The critical values $\kappa^l(p)$ and $\kappa^u(p)$ are computed via simulations as described below. An important feature of the CLR procedure is that the critical values $\kappa^l(p)$ and $\kappa^u(p)$ are computed by simulation not based on the indexing sets $\mathcal{V}^l$ and $\mathcal{V}^u$, but instead based on the preliminary set estimators $\widehat{V^l}$ and $\widehat{V^u}$ of, respectively:

$$V_0^l = \arg\max_{v \in \mathcal{V}^l} \{\theta^l(v)\}$$

and

$$V_0^u = \arg\min_{v \in \mathcal{V}^u} \{\theta^u(v)\}$$

Intuitively, $\widehat{V^l}$ (respectively, $\widehat{V^u}$) selects those bounding functions that are close enough to binding to affect the asymptotic distribution of the estimator of the lower bound $\widehat{\theta^l}(p)$ (upper bound $\widehat{\theta^u}(p)$). This is done because choosing the maximum or minimum over all possible bounding functions by using $\mathcal{V}^l$ and $\mathcal{V}^u$, respectively, leads to asymptotically valid but conservative inference. Below we describe how the preliminary set estimators $\widehat{V^l}$ and $\widehat{V^u}$—which CLR call adaptive inequality selectors—are computed.

First, consider the lower bound, and more specifically, computing $\kappa^l(p)$ and $\widehat{V^l}$. Let $\widehat{\boldsymbol{\gamma}^l}$ be the 16-dimensional column vector of sample analog estimators of all the unadjusted bounding functions for the lower bound, with $\widehat{\boldsymbol{\gamma}^u}$ defined likewise for the upper bounds. An initial step obtains from $B = 999$ bootstrap replications a consistent estimate $\widehat{\Omega}_l$ of the asymptotic variance-covariance matrix of $\sqrt{N}\left(\widehat{\boldsymbol{\gamma}^l} - \boldsymbol{\gamma}^l\right)$, where N denotes the sample size (an analogous process is followed for the upper bounds). With $\widehat{\boldsymbol{g}^l}(\boldsymbol{v})'$ the v^{th} row of $\widehat{\Omega}_l^{1/2}$, define $s^l(v) \equiv \frac{||\widehat{g^l}(v)||}{\sqrt{N}}$. Next, following CLR, we simulate $R = 100{,}000$ draws from a $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$ distribution, where $\boldsymbol{I}$ is the 16×16 identity matrix. The draws are labelled $\boldsymbol{Z}_r$, $r = 1, \ldots, 100{,}000$, and are used to compute $Z_r^*(v) \equiv \widehat{\boldsymbol{g}^l}(\boldsymbol{v})'\boldsymbol{Z}_r / \left|\left|\widehat{g^l}(v)\right|\right|$ for each r and v. In each replication, we select the maximum over the set of $Z_r^*(1), \ldots, Z_r^*(16)$. From the resulting R values, we compute $\kappa^l(c)$, which is defined as the c^{th} quantile of these values, where $c \equiv 1 - (0.1/\, log\, N)$. The value of $\kappa^l(c)$ is then used to construct the following set estimator:

$$\widehat{V^l} = \left\{v \in \mathcal{V}^l : \widehat{\theta^l}(v) \geq \max_{\tilde{v} \in \mathcal{V}^l}\left[\widehat{\theta^l}(\tilde{v}) - \kappa^l(c) \cdot s^l(\tilde{v})\right] - 2\kappa^l(c) \cdot s^l(v)\right\}$$

From the values $Z_r^*(v)$, we next take the maximum from each replication r, this time restricting the search only to $v \in \widehat{V^l}$ (instead of searching over all the indexes $v \in \mathcal{V}^l$). Lastly, the CLR critical value $\kappa^l(p)$ is taken as the p^{th} quantile of these R values (i.e., as the p^{th} quantile of the R maximums coming from each replication).

Regarding computation of $\kappa^u(p)$ and $\widehat{V^u}$ for the upper bound, the same procedure as above is followed, now defining $\widehat{V^u}$ as:[3]

$$\widehat{V^u} = \left\{v \in \mathcal{V}^u : \widehat{\theta^u}(v) \leq \min_{\tilde{v} \in \mathcal{V}^u} \left[\widehat{\theta^u}(\tilde{v}) + \kappa^u(c) \cdot s^u(\tilde{v})\right] + 2\kappa^u(c) \cdot s^u(v)\right\}.$$

Half-median unbiased estimators of the lower and upper bounds of the average treatment effect are obtained by setting $p = 0.5$, computing the critical values $\kappa^l(0.5)$ and $\kappa^u(0.5)$ as described above, and using equations (13) and (14) to compute the half-median unbiased estimates $\widehat{\theta^l}(0.5)$ and $\widehat{\theta^u}(0.5)$.

To obtain $(1-\alpha)\cdot 100\%$ confidence intervals for the true value of the average treatment effect θ_0, we must make one final adjustment which accounts for the width of the identified set. Borrowing notation from CLR (2013), define:

$$\hat{\Gamma} \equiv \widehat{\theta^u}(0.5) - \widehat{\theta^l}(0.5)$$

$$\widehat{\Gamma^+} \equiv \max\{0, \hat{\Gamma}\}$$

$$\rho \equiv \max\{\widehat{\theta^u}(0.75) - \widehat{\theta^u}(0.25)\,,\ \widehat{\theta^l}(0.25) - \widehat{\theta^l}(0.75)\}$$

$$\tau \equiv 1/(\rho \log N)$$

$$\hat{p} \equiv 1 - \Phi\left(\tau\widehat{\Gamma^+}\right) \cdot \alpha,$$

where $\Phi(\cdot)$ is the standard normal CDF. Note that $\hat{p} \in [1-\alpha, 1-\alpha/2]$, with $\hat{p}$ approaching $1-\alpha$ when $\hat{\Gamma}$ grows large relative to sampling error, and $\hat{p} = 1-\alpha/2$ when $\hat{\Gamma}$ =0. An asymptotically

[3] Because of the symmetry of the normal distribution, no changes are needed in any of the other steps.

valid $(1-\alpha)\cdot 100\%$ confidence interval for the true value of θ_0 is given by $\left[\widehat{\theta^l}(\hat{p}),\ \widehat{\theta^u}(\hat{p})\right]$. We report 95% confidence intervals for θ_0 using the critical values $\kappa^l(\hat{p})$ and $\kappa^u(\hat{p})$ with $\alpha = 0.05$ in equations (A13) and (A14), respectively.

References

Chernozhukov, V., Lee, S., & Rosen, A. M. (2009). Intersection bounds: estimation and inference. Center for Microdata Methods and Practice (cemmap) Working Paper CWP19/09.

Chernozhukov, V., Lee, S., & Rosen, A. M. (2013). Intersection bounds: estimation and inference. *Econometrica*, *81*(2), 667-737.

De Haan, M. (2011). The effect of parents' schooling on child's schooling: a nonparametric bounds analysis. *Journal of Labor Economics*, *29*(4), 859-892.

Flores, C. A., & Flores-Lagunes, A. (2013). Partial identification of local average treatment effects with an invalid instrument. *Journal of Business & Economic Statistics*, *31*(4), 534-545.

Hirano, K., & Porter, J. R. (2012). Impossibility results for nondifferentiable functionals. *Econometrica*, *80*(4), 1769-1790.

Imbens, G. W., & Manski, C. F. (2004). Confidence intervals for partially identified parameters. *Econometrica*, *72*(6), 1845-1857.

Stoye, J. (2009). More on confidence intervals for partially identified parameters. *Econometrica*, *77*(4), 1299-1315.

Tamer, E. (2010). Partial identification in econometrics. *Annual Review of Economics*, *2*, 167-195.

Manski, C. F. (1989). Anatomy of the selection problem. *Journal of Human Resources*, 343-360.

Manski, C. F. (1997). Monotone treatment response. *Econometrica*, 65, 1311-1334.

Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables with an application to the returns to schooling. *Econometrica,* 68(July), 997-1010

Appendix B: Description of Cognition Tests

Mini-Mental State Examination (MMSE): This test includes 22 items (10 orientation, 8 language, 1 of each: registration, memory, spelling backward, and construction). The maximum score is 30.

HRS Telephone Interview for Cognition Status (TICS): Participants had to identify two name two words (vocabulary) and naming the President of the United States. Specific questions were: (1) What do people usually use to cut paper? (2) What do you call the kind of prickly plant that grows in the desert? (3) Who is the President of the United States right now? The maximum score is 3.

Consortium to Establish a Registry for Alzheimer's Disease Word (CERAD) Word List Learning and Recall- Immediate: Participants were shown a list of 10 words, two seconds at a time for each word. Participants read each word and after the last word were asked to recall as many words from the list as possible. The score ranges from 0-10. Participants undertook three trails of the task.

Animal Naming: Participants were asked to name as many animals as they could within a 60-second time limit. The test score range is 0-43.

Letter Cancellation: Participants had one minute to cross out as many "P" and "W" letters as possible from a large grid of letters.

Backward Counting: Participants had to count backward from 100 as fast as possible in a 30 second time limit.

Community Screening Instrument for Dementia (CSI-D): Participants were asked questions evaluating language, knowledge and the ability to follow directions. The questions/tasks were: (1) point to your elbow; (2) what do you do with a hammer? (3) where is the local market/ local store?; (4) point first to the window and then to the door. The maximum score is 4.

CERAD Word List-Delayed: This is a single trial to recall the list of 10 words from the CERAD Word List Learning and Recall (Immediate) task. Participants are asked to freely recall as many words as possible from that list. The interviewer records all correct responses as well as intrusions (words not on the original list). Respondents are given up to 2 minutes to complete this task.

Story Memory-Immediate: Participants were read one of two from the Wechsler Memory Scale (WMS-IV). participants had to report back on the main parts of the story immediately after it was read.

CERAD Word List-Recognition: Participants were visually presented a series of 20 words, 10 from the CERAD word list and 10 foils. They were asked to identify which words were given on the original list

CERAD Constructional Praxis – Immediate: Participants had to copy geometric figures that varied in difficulty

Symbol-Digit Modalities Test (SDMT): Participants were given random geometric figures and a separate key that paired numbers with each figure. Participants had to substitute a number of each figure, completing as many pairings as possible in the 90-second time limit.

CERAD Constructional Praxis – Delayed: This is a delayed recall of the geometric shaped drawn in the test of CERAD Constructional Praxis – Immediate. Respondents are asked to draw the shapes from earlier in the interview to the best of their memory.

Story Memory-Delayed: Participants were asked to think back to the two stories read to them earlier and recall as much about each story as they can.

Story Memory-Recognition: Participants were given 15 yes/no questions on whether a specified story point was part of the story they were read

HRS Number Series: Participants were presented with a series of numbers with one or two numbers missing. Participants had to identify the missing numbers. The test was not timed and was adaptive such that difficulty level changed depending on the participants' responses. The range is 409-584.

Raven's Standard Progressive Matrices: This test evaluates picture-based pattern reasoning of varying difficulty. Each question presents a geometric picture with a small section that appears to have been cut out. Participants are shown a set of smaller pictures that fit the missing piece and are asked to identify which is the correct one to complete the pattern.

Trail Making Test (A and B): Participants asked to draw lines connecting consecutively numbered circles on a worksheet (part A) and connect consecutively numbered and lettered circles on another worksheet (part B) by alternating between the numbers and letters. The interviewer is instructed to point out errors to the participant and have the participant go back to the previous circle and move on to the next correct one. The score for this test is the number of seconds to complete part A and part B, where the time to correct errors serves to increase the total time to complete the test.

Appendix C: Additional Tables

Appendix Table C1: Literature Summary

Study	Country and Data		Method	Age	Sample	Outcome	β_{OLS} (95% CI)	B_{Causal} (95% CI)
Glymour et al. (2008)	US; HRS		IV based on changes in compulsory school leaving age over 1907-1962	Range: 55-101	Men+Women	Immediate+ Delayed Memory	0.09 (0.08, 0.10)	0.34 (0.11, 0.57)
					Men+Women	Mental Status	0.15 (0.14, 0.16)	-0.06 (-0.37, 0.26)
Herd & Sicinski (2023)	US; WLS		Sibling fixed-effects controlling for education polygenic score	Mean: 72	Men+Women	Delayed Memory	0.031 (0.01, 0.05)	0.032 (0.00, 0.06)
					Men+Women	Immediate Memory	0.054 (0.03, 0.07)	0.067 (0.03, 0.08)
Banks & Mazzonna (2012)	England; ELSA		IV based on 1947 Raising of School leaving age from 14 to 15	Range: 72-74	Men	Immediate+Delayed Memory	Not Reported	0.51 (0.07, 0.96)
					Women	Immediate+Delayed Memory	Not Reported	0.52 (-0.02, 1.06)
					Men	Executive Function	Not Reported	0.55 (0.12, 0.98)
					Women	Executive Function	Not Reported	0.20 (-0.39, 0.79)

Study	Country and Data		Method	Age	Sample	Outcome	β_{OLS} (95% CI)	B_{Causal} (95% CI)
Goorman (2023)	England; Understanding Society		IV based on 1972 Raising of School leaving age from 15 to 16	Range: 48-60	Men+Women	Immediate+Delayed Memory	0.24 (0.24, 0.24)	0.42 (0.05, 0.79)
						Verbal Fluency	0.23 (0.23, 0.23)	0.05 (-0.34, 0.44)
Fletcher et al. (2021)	UK; UK Biobank		Sibling fixed-effects with controls for Alzheimer's Disease, cognition and education polygenic scores	Mean: 57	Men+Women	Fluid Intelligence	0.21 (0.20, 0.23)	0.12 (0.09, 0.15)
Schneeweis et al. (2014)	European Countries; SHARE		IV based on education reforms	Mean age: 61	Men+Women	Immediate Memory	0.07 (0.06, 0.07)	0.08 (0.01, 0.16)
					Men+Women	Delayed Memory	0.06 (0.06, 0.07)	0.09 (0.01, 0.16)
							0.06 (0.06, 0.07)	-0.03 (-0.11, 0.05)
					Men+Women	Verbal Fluency	0.06 (0.06, 0.07)	-0.03 (-0.05, 0.11)
					Men+Women	Good Numeracy	0.03 (0.03, 0.03)	-0.01 (-0.06, 0.03)

Study	Country and Data		Method	Age	Sample	Outcome	β_{OLS} (95% CI)	B_{Causal} (95% CI)
					Men+Women	Good Orientation to Date	0.00 (0.00, 0.01)	-0.01 (-0.03, 0.02)

Notes: Estimates are taken from the following tables: Glymour et al. (2008) Table 3 model 4; Banks & Mazzonna (2012) Table 5 with a bandwidth of 2 years and conditional on leaving school before age 16; Fletcher et al. (2021) Table 2 column 3 (OLS) and column 6 (FE); Schneeweis et al. (2014) Table 5 baseline specification. Effect size: standard deviation per 1 additional grade of schooling, apart from good numeracy and good orientation in Schneeweis et al. (2014) which are dummy variables. Fletcher et al. (2021) measure schooling with a series of dummy variable for years of schooling and the outcome is not standardized. We report estimates on the dummy variable years of education=10. We divide the estimates by 3 since the reference category is 7 years of education, and then divide by the SD of the outcome (2.1) We divide the estimates by 6 (we assume a difference of 6 grades between individuals with primary schooling and those without) and then by the standard deviation of the outcome. HRS: Health & Retirement Study. ELSA: English Longitudinal Study of Ageing. WLS: Wisconsin Longitudinal Study. SHARE: Survey of Health, Ageing, and Retirement in Europe.

Table C2: Cognition Tests in HCAP

Test	Domains Assessed	Observations	Number of Imputed Observations
Mini Mental State Examination	O; MemI; MemD; VF; AS; Viz	3,347	0
HRS TICS	MemI	3,344	3
CERAD Word List-Immediate Trail 1	MemI	3,338	9
CERAD Word List-Immediate Trail 2	MemI	3,330	17
CERAD Word List-Immediate Trail 3	MemI	3,323	24
Animal Naming	VF	3,345	2
Letter Cancellation	AS	3,197	150
Backward Counting	AS	3,302	45
Community Screening Instrument for Dementia (CSI-D)	O; EF; VF	3,341	6
CERAD Word List-Delayed	MemD	3,341	16
Story Recall Immediate	MemI	3,306	41
CERAD Word List Recognition	MemR	3,323	24
CERAD Constructional Praxis-Immediate	Viz	3,308	39
Symbol Digit Modalities Test (SDMT)	EF; AS	3,168	179
CERAD Constructional Praxis-Delayed	MemD	3,305	42
Story Recall -Delayed	MemD	3,316	31
Story Recall -Recognition	MemR	3,239	108
HRS Number Series	EF	2,769	578
Raven's matrices	EF	3,287	60
Trail making Part A	AS	3,242	105
Trail making Part B	AS	3,131	216

Notes: MemD: Delayed Memory MemI: Immediate Memory MemR: Recognition Memory O: Orientation VF: Verbal Fluency AS: Attention/Speed EF: Executive Function Viz: Visuospatial. Information taken from: https://hrs.isr.umich.edu/sites/default/files/meta/hcap/2016/codebook/hc16hp_ri.htm

Table C3: Average Cognition Scores by Schooling Attainment

	Mini Mental State Exam	Immediate + Delayed Memory	Recognition Memory	Verbal Fluency	Attention/ Speed	Executive Function	Visuospatial
A: Full Sample							
HS Dropout	24.28 (4.56)	37.55 (15.95)	26.64 (4.51)	12.76 (5.33)	32.42 (14.37)	9.54 (3.87)	6.88 (2.39)
HS Grads	26.71 (3.60)	46.82 (17.22)	28.78 (4.41)	15.30 (6.04)	43.63 (13.56)	12.05 (3.57)	8.04 (2.24)
Some College	27.49 (3.46)	51.73 (16.55)	29.37 (4.09)	16.74 (6.29)	46.65 (13.52)	13.20 (3.09)	8.40 (2.15)
College Grad	27.87 (3.06)	56.46 (17.65)	29.87 (4.20)	19.07 (7.17)	50.12 (14.23)	14.22 (2.99)	9.16 (2.09)
B: Men							
HS Dropout	23.79 (4.51)	34.80 (15.34)	26.20 (4.33)	13.10 (5.69)	32.70 (13.19)	9.72 (3.99)	7.10 (2.21)
HS Grads	26.21 (3.65)	43.00 (15.83)	28.34 (4.20)	15.35 (6.09)	43.01 (13.81)	12.12 (3.66)	8.08 (2.28)
Some College	27.38 (3.27)	49.22 (14.49)	28.97 (4.09)	16.78 (6.12)	47.51 (13.26)	13.60 (3.10)	8.57 (2.14)
College Grad	27.83 (2.39)	53.71 (16.21)	29.66 (3.54)	19.02 (7.06)	51.19 (13.58)	14.57 (2.81)	9.31 (1.91)
C: Women							
HS Dropout	24.54 (4.57)	38.98 (16.10)	26.88 (4.60)	12.58 (5.13)	32.28 (14.96)	9.45 (3.81)	6.77 (2.48)
HS Grads	27.01 (3.54)	49.07 (17.61)	29.05 (4.51)	15.27 (6.02)	43.99 (13.41)	12.01 (3.52)	8.01 (2.21)
Some College	27.56 (3.57)	53.20 (17.50)	29.60 (4.07)	16.71 (6.39)	46.15 (13.65)	12.97 (3.05)	8.30 (2.15)
College Grad	27.91 (3.56)	58.96 (18.53)	30.06 (4.72)	19.12 (7.27)	49.15 (14.73)	13.91 (3.12)	9.01 (2.23)
D: Non-Hispanic White							
HS Dropout	25.28 (4.17)	39.52 (16.82)	27.35 (4.45)	12.98 (5.70)	35.93 (12.95)	10.47 (3.72)	7.28 (2.38)
HS Grads	27.00 (3.59)	47.66 (17.29)	28.88 (4.44)	15.71 (5.96)	45.21 (12.70)	12.57 (3.28)	8.24 (2.18)
Some College	27.66 (3.20)	52.47 (16.52)	29.58 (3.93)	17.21 (6.19)	47.39 (13.58)	13.65 (2.90)	8.58 (2.05)
College Grad	28.01 (2.93)	57.18 (17.54)	30.13 (3.96)	19.64 (7.13)	50.87 (13.89)	14.58 (2.71)	9.29 (2.04)
E: Non-Hispanic Black							
HS Dropout	23.21 (4.45)	35.75 (15.21)	26.66 (4.33)	11.55 (4.85)	28.96 (13.51)	8.12 (3.93)	6.42 (2.23)
HS Grads	25.28 (3.49)	42.38 (16.60)	28.59 (4.50)	12.75 (5.48)	36.04 (15.60)	9.18 (3.86)	6.87 (2.28)
Some College	26.80 (4.08)	48.78 (17.09)	28.63 (5.07)	14.59 (6.71)	45.56 (12.80)	11.37 (3.28)	7.48 (2.30)
College Grad	27.08 (3.96)	53.06 (19.07)	28.59 (5.65)	16.17 (7.05)	45.09 (14.86)	12.05 (3.62)	8.14 (2.35)
F: Hispanic							
HS Dropout	23.67 (4.87)	36.29 (14.86)	25.70 (4.48	13.30 (4.89)	29.89 (15.77)	9.18 (3.64)	6.57 (2.46)
HS Grads	27.06 (2.68)	47.68 (16.67)	28.21 (3.91)	16.60 (6.77)	42.59 (11.85)	12.49 (2.96)	8.19 (1.93)
Some College	27.12 (4.48)	49.41 (15.55)	28.53 (3.55)	15.86 (5.86)	43.35 (12.51)	12.37 (2.97)	8.43 (2.43)
College Grad	27.96 (2.11)	52.44 (13.01)	28.32 (3.48)	17.20 (5.94)	48.04 (16.21)	13.24 (4.01)	9.12 (2.02)

Notes: Standard deviations in parentheses.

Table C4: MTS+MIV and MTR+MIV Bounds for the Effect of Schooling on Immediate and Delayed Memory

	MTS+MIV	**MTR+MIV**
HS Grad vs HS Dropout	[-48.861, 33.012] (-49.640, 34.976)	[0.000, 45.785] (0.000, 46.864)
Some College vs HS Grad	[-46.460, 25.190] (-47.081, 26.387)	[5.977, 15.253] (4.549, 17.012)
College vs Some College	[-49.038, 30.043] (-49.757, 31.357)	[17.255, 13.562] (15.130, 16.105)
College vs HS Grad	[-57.499, 17.081] (-58.279, 18.512)	[25.239, 26.952] (23.826, 29.126)
College vs HS Dropout	[-72.891, 16.299] (-73.776, 18.357)	[32.543, 65.050] (31.530, 67.068)

Notes: Estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome was used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school.

Table C5: Bounds on Expected Immediate and Delayed Memory Scores for Levels of Schooling Attainment

Assumptions	High-school Dropout	High-school Graduate	Some College	College Graduate
No Assumption	[6.684, 87.238] (6.221, 87.929)	[15.438, 81.127] (14.756, 81.885)	[12.054, 87.217] (11.365, 87.847)	[14.627, 87.239] (13.860, 87.836)
MTS	[37.545, 87.238] (36.436, 87.929)	[38.483, 72.015] (37.604, 72.876)	[25.457, 63.721] (24.567, 64.683)	[14.627, 56.464] (13.860, 57.465)
MTR	[6.684, 48.814] (6.221, 49.321)	[22.124, 59.578] (21.407, 60.273)	[34.180, 76.454] (33.452, 77.163)	[48.811, 87.239] (48.303, 87.836)
MTS+MTR	[37.545, 48.814] (36.417, 49.330)	[45.169, 50.465] (44.384, 51.049)	[47.584, 52.957] (46.942, 53.789)	[48.811, 56.464] (48.284, 57.502)
MTS+MIV	[39.207, 87.310] (37.413, 87.990)	[38.428, 72.000] (37.517, 72.834)	[25.425, 63.631] (24.573, 64.527)	[14.417, 55.471] (13.698, 56.533)
MTR+MIV	[10.640, 48.909] (9.869, 49.394)	[48.690, 56.348] (48.198, 57.148)	[62.213, 63.902] (60.969, 65.870)	[81.723, 75.591] (80.578, 77.816)
MTS+MTR+MIV	[39.207, 48.909] (37.294, 49.425)	[48.690, 50.510] (48.161, 51.062)	[50.166, 52.892] (49.575, 53.681)	[52.352, 55.471] (51.452, 56.655)

Notes: Estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome were used in computing the bounds.The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school. The estimated bounds on $E[Y(t)]$ presented in this table provide a more nuanced picture of the estimated bounds in Table 2. It is important to note, however, that the estimated bounds on the population ATE in Table 2 cannot be directly obtained from the bounds in Table C4 by using equation (3) in the main text. Specifically, while close, the estimated lower (respectively, upper) bounds in Table 2 will not in general be numerically equal to the difference between the estimated lower (upper) and upper (lower) bounds of the corresponding average potential outcomes $E[Y(t)]$. For example, under the MTS+MTR+MIV assumptions, the estimated lower bound on the effect from secondary (t_2) to tertiary (t_4) in Table 2 of 1.87, is close but not numerically equal to the difference between the lower bound on $E[Y(t_4)]$ in Table C4 of 52.35 and the upper bound on $E[Y(t_2)]$ of 50.51, which equals 1.84. The reason is that, contrary to expectations, the Chernozhukov et al. (2013) bias-correction method employed to estimate the bounds does not have the property that the estimates from the method applied to the bounds on the effect (e.g., $\Delta(t_2, t_4) = E[Y(t_4) - Y(t_2)]$) equal the difference in the corresponding estimates from separately applying the method to the bounds of each of the mean potential outcomes (e.g., $E[Y(t_2)]$ and $E[Y(t_4)]$). See appendix A for further details on the Chernozhukov et al. (2013) procedure.

Table C6: OLS Estimates and Bounds for the Effect of Schooling on Other Cognition Domains

	Control Mean (SD)	OLS	MTS+MTR	MTS+MTR+MIV
	(1)	(2)	(3)	(4)
Panel A: Mini-Mental State Examination				
HS Grad vs HS Dropout	24.28 (4.56)	2.433*** (0.225)	[0.000, 2.916] (0.000, 3.260)	[0.000, 2.916] (0.000, 3.318)
Some College vs HS Grad	26.71 (3.60)	0.780*** (0.172)	[0.000, 1.312] (0.000, 1.554)	[0.000, 0.864] (0.000, 1.002)
College Grad vs Some College	27.49 (3.46)	0.379** (0.167)	[0.000, 1.209] (0.000, 1.451)	[0.000, 0.537] (0.000, 0.753)
College Grad vs HS Grad	26.71 (3.60)	1.159*** (0.157)	[0.000, 1.593] (0.000, 1.849)	[0.289, 0.954] (0.122, 1.164)
College Grad vs HS Dropout	24.28 (4.56)	3.592*** (0.223)	[0.000, 3.593] (0.000, 3.969)	[0.710, 3.369] (0.557, 3.818)
Panel B: Recognition Memory				
HS Grad vs HS Dropout	26.64 (4.51)	2.141*** (0.238)	[0.000, 2.559] (0.000, 2.914)	[0.000, 2.318] (0.000, 3.018)
Some College vs HS Grad	28.78 (4.41)	0.581*** (0.206)	[0.000, 1.094] (0.000, 1.395)	[0.000, 0.725] (0.000, 0.887)
College Grad vs Some College	29.37 (4.09)	0.505** (0.213)	[0.000, 1.182] (0.000, 1.494)	[0.000, 0.535] (0.000, 0.832)
College Grad vs HS Grad	28.78 (4.41)	1.086*** (0.203)	[0.000, 1.468] (0.000, 1.808)	[0.197, 0.885] (0.003, 1.169)
College Grad vs HS Dropout	26.64 (4.51)	3.227*** (0.244)	[0.000, 3.228] (0.000, 3.641)	[0.552, 2.751] (0.375, 3.503)
Panel C: Verbal Fluency				
HS Grad vs HS Dropout	12.76 (5.33)	2.540*** (0.297)	[0.000, 3.857] (0.000, 4.294)	[0.000, 3.581] (0.000, 4.516)
Some College vs HS Grad	15.30 (6.04)	1.441*** (0.302)	[0.000, 2.500] (0.000, 2.923)	[0.000, 1.266] (0.000, 1.493)
College Grad vs Some College	16.74 (6.29)	2.338*** (0.346)	[0.000, 3.524] (0.000, 3.998)	[0.000, 2.026] (0.000, 2.406)
College Grad vs HS Grad	15.30 (6.04)	3.779*** (0.317)	[0.000, 4.233] (0.000, 4.743)	[0.627, 2.403] (0.352, 2.773)
College Grad vs HS Dropout	12.76 (5.33)	6.319*** (0.341)	[0.000, 6.321] (0.000, 6.889)	[1.019, 5.461] (0.765, 6.492)
Panel D: Executive Function				
HS Grad vs HS Dropout	9.54 (3.87)	2.509*** (0.200)	[0.000, 3.342] (0.000, 3.641)	[0.000, 2.831] (0.000, 3.194)
Some College vs HS Grad	12.05 (3.57)	1.154*** (0.161)	[0.000, 1.865] (0.000, 2.087)	[0.000, 0.973] (0.000, 1.097)
College Grad vs Some College	13.20 (3.09)	1.020*** (0.157)	[0.000, 2.053] (0.000, 2.261)	[0.000, 1.123] (0.000, 1.314)
College Grad vs HS Grad	12.05 (3.57)	2.173*** (0.154)	[0.000, 2.621] (0.000, 2.859)	[0.454, 1.522] (0.311, 1.703)
College Grad vs HS Dropout	9.54 (3.87)	4.682*** (0.197)	[0.000, 4.684] (0.000, 5.006)	[0.872, 3.887] (0.736, 4.285)

Panel E: Attention/Speed				
HS Grad vs HS Dropout	32.42 (14.37)	11.207*** (0.748)	[0.000, 13.598] (0.000, 14.686)	[0.000, 12.153] (0.000, 13.491)
Some College vs HS Grad	43.63 (13.56)	3.023*** (0.601)	[0.000, 5.921] (0.000, 6.866)	[0.000, 3.621] (0.000, 4.159)
College Grad vs Some College	46.56 (13.52)	3.471*** (0.714)	[0.000, 7.005] (0.000, 8.046)	[0.000, 3.258] (0.000, 4.157)
College Grad vs HS Grad	43.63 (13.56)	6.494*** (0.660)	[0.000, 8.493] (0.000, 9.589)	[1.253, 5.025] (0.632, 5.900)
College Grad vs HS Dropout	32.42 (14.37)	17.701*** (0.795)	[0.000, 17.706] (0.000, 19.009)	[3.097, 15.108] (2.516, 16.683)
Panel F: Visuospatial				
HS Grad vs HS Dropout	6.88 (2.39)	1.156*** (0.124)	[0.000, 1.532] (0.000, 1.722)	[0.000, 1.521] (0.000, 1.755)
Some College vs HS Grad	8.04 (2.24)	0.366*** (0.101)	[0.000, 0.768] (0.000, 0.914)	[0.000, 0.395] (0.000, 0.479)
College Grad vs Some College	8.40 (2.15)	0.753*** (0.109)	[0.000, 1.146] (0.000, 1.293)	[0.000, 0.710] (0.000, 0.836)
College Grad vs HS Grad	8.04 (2.24)	1.120*** (0.102)	[0.000, 1.326] (0.000, 1.486)	[0.103, 0.905] (0.011, 1.025)
College Grad vs HS Dropout	6.88 (2.39)	2.276*** (0.126)	[0.000, 2.277] (0.000, 2.488)	[0.301, 2.176] (0.214, 2.433)

Notes: Robust standard errors in (.) in column 2. In columns 3 and 4 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome was used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school.

Table C7: Bounds on Effects of Schooling on Cognition with Alternative Education Coding Schemes

		Mini-Mental State Exam	Immediate +Delayed Memory	Recognition Memory	Attention/ Speed	Verbal Fluency	Executive Function	Visuospatial
Grades of Schooling								
Treatment	**Control**							
10	≤9	[0.000, 3.652] (0.000, 4.048)	[0.008, 13.969] (0.000, 17.486)	[0.000, 3.435] (0.000, 3.917)	[0.000, 15.417] (0.000, 17.168)	[0.000, 3.996] (0.000, 4.916)	[0.000, 3.623] (0.000, 4.147)	[0.000, 1.830] (0.000, 2.106)
11	10	[0.000, 0.323] (0.000, 0.425)	[0.000, 1.327] (0.000, 1.716)	[0.106, 0.312] (0.000, 0.427)	[0.000, 1.177] (0.000, 1.485)	[0.000, 0.272] (0.000, 0.416)	[0.000, 0.219] (0.000, 0.310)	[0.000, 0.140] (0.000, 0.191)
12	11	[0.000, 0.275] (0.000, 0.396)	[0.000, 1.142] (0.000, 1.820)	[0.000, 0.087] (0.000, 0.318)	[0.000, 1.361] (0.000, 1.690)	[0.000, 0.732] (0.000, 0.905)	[0.000, 0.269] (0.000, 0.364)	[0.000, 0.189] (0.000, 0.284)
13	12	[0.000, 0.637] (0.000, 0.883)	[0.000, 2.719] (0.000, 3.820)	[0.000, 0.529] (0.000, 0.787)	[0.000, 2.763] (0.000, 3.668)	[0.000, 0.863] (0.000, 1.261)	[0.000, 0.721] (0.000, 0.947)	[0.000, 0.242] (0.000, 0.377)
14	13	[0.000, 0.475] (0.000, 0.717)	[0.000, 2.974] (0.000, 4.047)	[0.000, 0.439] (0.000, 0.738)	[0.000, 1.697] (0.000, 2.552)	[0.000, 1.116] (0.000, 1.519)	[0.000, 0.571] (0.000, 0.778)	[0.000, 0.194] (0.000, 0.345)
15	14	[0.000, 0.239] (0.000, 0.751)	[0.000, 2.176] (0.000, 4.134)	[0.000, 0.257] (0.000, 0.785)	[0.000, 2.162] (0.000, 4.246)	[0.000, 1.075] (0.000, 2.001)	[0.000, 0.823] (0.000, 1.235)	[0.000, 0.385] (0.000, 0.676)
16	15	[0.000, 0.369] (0.000, 0.707)	[0.000, 2.003] (0.000, 3.659)	[0.000, 0.370] (0.000, 0.854)	[0.000, 2.119] (0.000, 3.642)	[0.000, 0.779] (0.000, 1.481)	[0.000, 0.519] (0.000, 0.853)	[0.000, 0.568] (0.000, 0.787)
17	16	[0.000, 0.152] (0.000, 0.752)	[0.000, 4.764] (0.000, 7.306)	[0.000, 0.472] (0.000, 1.082)	[0.000, 1.466] (0.000, 4.724)	[0.000, 1.787] (0.000, 3.326)	[0.000, 0.568] (0.000, 1.033)	[0.000, 0.631] (0.000, 1.020)
16	12	[0.375, 0.850] (0.000, 1.094)	[2.338, 4.675] (0.344, 6.034)	[0.188, 0.731] (0.000, 1.078)	[3.229, 4.036] (0.292, 5.269)	[1.062, 1.811] (0.000, 2.341)	[0.889, 1.191] (0.483, 1.484)	[0.230, 0.763] (0.000, 0.934)

Notes: Robust standard errors in (.) in column 2. In columns 3-7 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the outcome was used in computing the bounds. The MIV is mother's grades of schooling divided into 3 bin

Table C8: Inverse Probability Weighted OLS Estimates and Bounds for the Effect of Schooling on Cognition

	Control Mean (SD)	OLS	MTS+MTR	MTS+MTR+MIV
	(1)	(2)	(3)	(4)
Panel A: Immediate+Delayed Memory				
HS Grad vs HS Dropout	37.55 (15.95)	10.477*** (0.202)	[0.000, 13.701] (0.000, 16.692)	[0.000, 12.657] (0.000, 15.606)
Some College vs HS Grad	46.82 (17.22)	6.488*** (1.365)	[0.000, 9.439] (0.000, 11.499)	[0.000, 6.231] (0.000, 7.601)
College Grad vs Some College	51.73 (16.55)	1.928 (1.667)	[0.000, 8.153] (0.000, 11.040)	[0.000, 4.239] (0.000, 6.687)
College Grad vs HS Grad	46.82 (17.22)	8.416*** (1.887)	[0.000, 10.922] (0.000, 13.986)	[2.654, 6.506] (1.164, 8.930)
College Grad vs HS Dropout	37.55 (15.95)	18.892*** (2.234)	[0.000, 18.903] (0.000, 22.602)	[5.082, 16.157] (3.647, 19.853)
Panel B: Mini-Mental State Examination				
HS Grad vs HS Dropout	24.28 (4.56)	3.627*** (0.734)	[0.000, 3.638] (0.000, 4.792)	[0.000, 3.457] (0.000, 4.370)
Some College vs HS Grad	26.71 (3.60)	1.270*** (0.396)	[0.000, 2.059] (0.000, 2.675)	[0.000, 1.543] (0.000, 1.990)
College Grad vs Some College	27.49 (3.46)	0.283 (0.259)	[0.000, 1.734] (0.000, 2.232)	[0.000, 0.846] (0.000, 1.322)
College Grad vs HS Grad	26.71 (3.60)	1.553*** (0.377)	[0.000, 2.276] (0.000, 2.889)	[0.509, 1.528] (0.058, 2.012)
College Grad vs HS Dropout	24.28 (4.56)	4.580*** (0.700)	[0.000, 4.583] (0.000, 5.722)	[1.186, 4.124] (0.748, 5.053)
Panel C: Recognition Memory				
HS Grad vs HS Dropout	26.64 (4.51)	2.687*** (0.451)	[0.000, 2.991] (0.000, 3.728)	[0.000, 2.778] (0.000, 3.488)
Some College vs HS Grad	28.78 (4.41)	0.738** (0.285)	[0.000, 1.366] (0.000, 1.868)	[0.000, 1.166] (0.000, 1.534)
College Grad vs Some College	29.37 (4.09)	-0.059 (0.610)	[0.000, 1.008] (0.000, 2.084)	[0.000, 0.436] (0.000, 1.410)
College Grad vs HS Grad	28.78 (4.41)	0.679 (0.609)	[0.000, 1.322] (0.000, 2.408)	[0.258, 1.063] (0.000, 2.171)
College Grad vs HS Dropout	26.64 (4.51)	3.366*** (0.703)	[0.000, 3.370] (0.000, 4.572)	[0.840, 3.034] (0.500, 4.382)
Panel D: Verbal Fluency				
HS Grad vs HS Dropout	12.76 (5.33)	2.655*** (0.587)	[0.000, 3.829] (0.000, 4.688)	[0.000, 3.631] (0.000, 4.581)
Some College vs HS Grad	15.30 (6.04)	1.947*** (0.435)	[0.000, 2.923] (0.000, 3.568)	[0.000, 1.849] (0.000, 2.286)
College Grad vs Some College	16.74 (6.29)	1.469*** (0.449)	[0.000, 3.220] (0.000, 3.937)	[0.000, 1.801] (0.000, 2.336)
College Grad vs HS Grad	15.30 (6.04)	3.415*** (0.500)	[0.000, 4.050] (0.000, 4.834)	[0.927, 2.326] (0.433, 2.855)
College Grad vs HS Dropout	12.76 (5.33)	6.070*** (0.597)	[0.000, 6.073] (0.000, 7.065)	[1.467, 5.219] (1.007, 6.231)

Panel E: Executive Function				
HS Grad vs HS Dropout	9.54 (3.87)	2.559*** (0.478)	[0.000, 3.463] (0.000, 4.112)	[0.000, 2.942] (0.000, 3.478)
Some College vs HS Grad	12.05 (3.57)	1.626*** (0.374)	[0.000, 2.446] (0.000, 2.967)	[0.000, 1.502] (0.000, 1.827)
College Grad vs Some College	13.20 (3.09)	0.899*** (0.214)	[0.000, 2.442] (0.000, 2.837)	[0.000, 1.523] (0.000, 1.988)
College Grad vs HS Grad	12.05 (3.57)	2.525*** (0.382)	[0.000, 3.136] (0.000, 3.673)	[0.695, 2.057] (0.295, 2.449)
College Grad vs HS Dropout	9.54 (3.87)	5.083*** (0.367)	[0.000, 5.085] (0.000, 5.698)	[1.327, 4.297] (0.987, 4.864)
Panel F: Attention/Speed				
HS Grad vs HS Dropout	32.42 (14.37)	13.025*** (2.008)	[0.000, 15.078] (0.000, 18.184)	[0.000, 13.054] (0.000, 15.511)
Some College vs HS Grad	43.63 (13.56)	4.122*** (1.141)	[0.000, 7.517] (0.000, 9.412)	[0.000, 5.563] (0.000, 6.934)
College Grad vs Some College	46.56 (13.52)	1.226 (1.262)	[0.000, 6.701] (0.000, 9.022)	[0.000, 2.392] (0.000, 4.256)
College Grad vs HS Grad	43.63 (13.56)	5.348*** (1.490)	[0.000, 8.460] (0.000, 11.005)	[1.714, 5.193] (0.491, 7.135)
College Grad vs HS Dropout	32.42 (14.37)	18.373*** (2.079)	[0.000, 18.383] (0.000, 21.862)	[4.845, 14.889] (3.395, 17.772)
Panel G: Visuospatial				
HS Grad vs HS Dropout	6.88 (2.39)	1.088*** (0.244)	[0.000, 1.511] (0.000, 1.869)	[0.000, 1.464] (0.000, 1.803)
Some College vs HS Grad	8.04 (2.24)	0.637*** (0.193)	[0.000, 1.047] (0.000, 1.327)	[0.000, 0.614] (0.000, 0.785)
College Grad vs Some College	8.40 (2.15)	0.647*** (0.199)	[0.000, 1.272] (0.000, 1.599)	[0.000, 0.939] (0.000, 1.296)
College Grad vs HS Grad	8.04 (2.24)	1.283*** (0.242)	[0.000, 1.544] (0.000, 1.917)	[0.206, 1.179] (0.000, 1.504)
College Grad vs HS Dropout	6.88 (2.39)	2.372*** (0.249)	[0.000, 2.373] (0.000, 2.788)	[0.459, 2.314] (0.290, 2.781)

Notes: Robust standard errors in (.) in column 2. In columns 3 and 4 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the residuals were used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school

Table C9: Summary Statistics for older adults in the MIDUS

	Mean (SD)	Min (Max)	Observations
	(1)	(3)	(5)
Demographics			
Age	72.04 (5.15)	65 (84)	1,016
Female	0.56 (0.50)	0 (1)	1,016
White	0.95 (0.21)	0 (1)	1,016
White Missing	0.04 (0.18)	0 (1)	1,106
Mother: HS Dropout	0.52 (0.50)	0 (1)	1,016
Mother: HS Grad	0.28 (0.45)	0 (1)	1,016
Mother: More than HS Grad	0.19 (0.39)	0 (1)	1,016
Schooling			
Grades of Schooling	13.79 (2.69)	6 (20)	1,016
HS Dropout	0.10 (0.30)	0 (1)	1,016
HS Grad	0.31 (0.46)	0 (1)	1,016
Some College	0.29 (0.45)	0 (1)	1,016
College Grad	0.30 (0.45)	0 (1)	1,016
Cognition			
Immediate+Delayed Memory	9.07 (4.56)	0 (28)	950
Attention/Speed	30.78 (9.23)	0 (100)	1,010
Verbal Fluency	16.24 (5.54)	0 (38)	1,015

Notes: Standard deviations in parentheses. Analysis limited to individuals aged 65 years or older in the MIDUS Cognition Project.

Table C10: OLS Estimates and Bounds for the Effect of Schooling on Cognitive Abilities of Older Individuals in the MIDUS

	Control Mean (SD)	OLS	MTS+MTR	MTS+MTR+MIV
	(1)	(2)	(3)	(4)
Panel A: Immediate+Delayed Memory				
HS Grad vs HS Dropout	7.15 (3.72)	1.280*** (0.462)	[0.000, 2.046] (0.000, 2.742)	[0.000, 2.247] (0.000, 3.013)
Some College vs HS Grad	8.43 (4.18)	1.203*** (0.383)	[0.000, 1.376] (0.000, 1.928)	[0.000, 0.756] (0.000, 1.070)
College vs Some College	9.63 (4.92)	0.170 (0.401)	[0.000, 0.781] (0.000, 1.359)	[0.000, 0.557] (0.000, 1.186)
College Grad vs HS Grad	8.43 (4.18)	1.372*** (0.363)	[0.000, 1.494] (0.000, 2.103)	[0.408, 0.726] (0.016, 1.332)
College Grad vs HS Dropout	7.15 (3.72)	2.653*** (0.477)	[0.000, 2.655] (0.000, 2.452)	[0.507, 2.753] (0.175, 3.648)
Panel B: Verbal Fluency				
HS Grad vs HS Dropout	13.91 (5.04)	1.345** (0.580)	[0.000, 2.465] (0.000, 3.399)	[0.000, 2.784] (0.000, 3.956)
Some College vs HS Grad	15.26 (5.20)	0.682 (0.424)	[0.000, 1.533] (0.000, 2.150)	[0.000, 0.487] (0.000, 0.767)
College vs Some College	15.94 (5.25)	2.379*** (0.447)	[0.000, 2.796] (0.000, 3.444)	[0.000, 2.033] (0.000, 2.551)
College Grad vs HS Grad	15.26 (5.20)	3.062*** (0.438)	[0.000, 3.198] (0.000, 3.918)	[0.077, 2.158] (0.000, 2.683)
College Grad vs HS Dropout	13.91 (5.04)	4.407*** (0.597)	[0.000, 4.410] (0.000, 5.459)	[0.217, 4.570] (0.000, 5.925)
Panel C: Attention/Speed				
HS Grad vs HS Dropout	26.07 (8.45)	3.764*** (0.973)	[0.000, 5.079] (0.000, 6.754)	[0.000, 5.221] (0.000, 7.699)
Some College vs HS Grad	29.83 (8.50)	0.308 (0.695)	[0.000, 1.809] (0.000, 2.881)	[0.000, 0.826] (0.000, 1.446)
College vs Some College	30.14 (8.63)	3.761*** (0.756)	[0.000, 4.260] (0.000, 5.393)	[0.000, 2.582] (0.000, 3.421)
College Grad vs HS Grad	29.83 (8.50)	4.069*** (0.739)	[0.000, 4.442] (0.000, 5.684)	[0.003, 2.964] (0.000, 3.836)
College Grad vs HS Dropout	26.07 (8.45)	7.834*** (1.017)	[0.000, 7.839] (0.000, 9.674)	[0.366, 7.525] (0.000, 10.074)

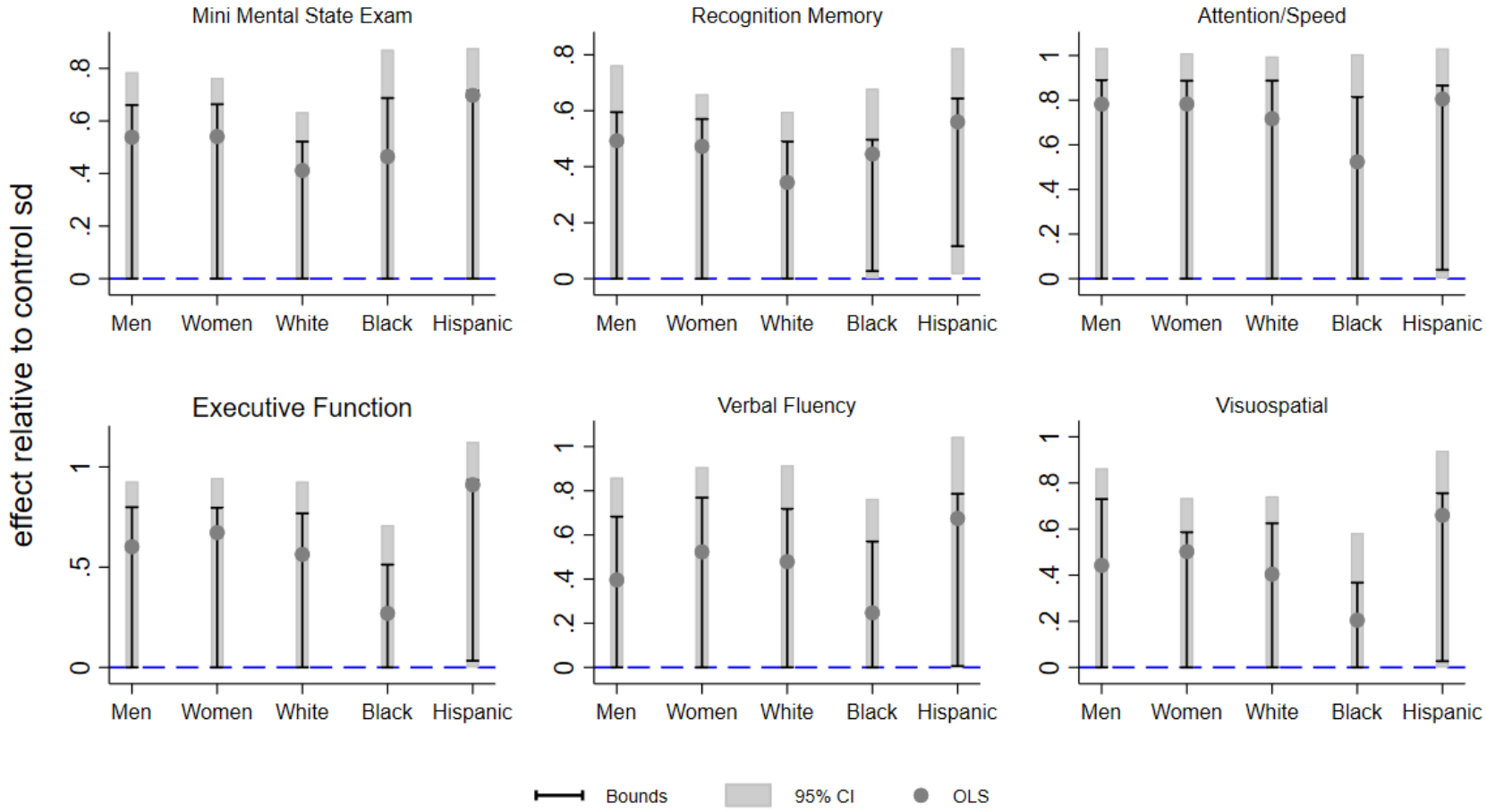
Notes: Robust standard errors in (.) in column 2. In columns 3 and 4 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the residuals were used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school. Analysis limited to individuals aged 65 years and over in the MIDUS Cognition Project. Estimation sample size in panels A ,B and C are 950, 1015, and 1010 observations respectively.

Table C11: OLS Estimates and Bounds for the Effect of Schooling on Cognitive Abilities of Middle-Aged Individuals in the MIDUS

	Control Mean (SD)	OLS	MTS+MTR	MTS+MTR+MIV
	(1)	(2)	(3)	(4)
Panel A: Immediate+Delayed Memory				
HS Grad vs HS Dropout	11.40 (4.64)	0.480 (0.627)	[0.000, 1.262] (0.000, 2.372)	[0.000, 1.067] (0.000, 2.490)
Some College vs HS Grad	11.88 (4.15)	0.655** (0.316)	[0.000, 0.987] (0.000, 1.466)	[0.000, 0.288] (0.000, 0.538)
College vs Some College	12.53 (4.41)	0.774*** (0.278)	[0.000, 0.982] (0.000, 1.430)	[0.000, 0.724] (0.000, 1.030)
College Grad vs HS Grad	11.88 (4.15)	1.429*** (0.295)	[0.000, 1.453] (0.000, 1.971)	[0.125, 0.749] (0.000, 1.073)
College Grad vs HS Dropout	11.40 (4.64)	1.909*** (0.608)	[0.000, 1.912] (0.000, 3.055)	[0.155, 1.681] (0.019, 3.125)
Panel B: Verbal Fluency				
HS Grad vs HS Dropout	17.89 (5.22)	0.427 (0.707)	[0.000, 2.738] (0.000, 3.899)	[0.000, 2.687] (0.000, 3.887)
Some College vs HS Grad	18.32 (5.29)	1.575*** (0.400)	[0.000, 2.776] (0.000, 3.363)	[0.000, 0.564] (0.000, 0.738)
College vs Some College	19.89 (5.81)	2.959*** (0.376)	[0.000, 3.427] (0.000, 3.983)	[0.000, 2.071] (0.000, 2.435)
College Grad vs HS Grad	18.32 (5.29)	4.534*** (0.387)	[0.000, 4.556] (0.000, 5.201)	[0.318, 2.087] (0.075, 2.453)
College Grad vs HS Dropout	17.89 (5.22)	4.961*** (0.693)	[0.000, 4.964] (0.000, 6.188)	[0.337, 4.602] (0.093, 5.837)
Panel C: Attention/Speed				
HS Grad vs HS Dropout	37.32 (12.54)	1.769 (1.649)	[0.000, 5.031] (0.000, 7.998)	[0.000, 6.276] (0.000, 9.678)
Some College vs HS Grad	39.09 (10.24)	2.674*** (0.767)	[0.000, 4.090] (0.000, 5.286)	[0.000, 1.010] (0.000, 1.355)
College vs Some College	41.77 (10.90)	3.325*** (0.696)	[0.000, 4.162] (0.000, 5.203)	[0.000, 2.875] (0.000, 3.554)
College Grad vs HS Grad	39.09 (10.24)	6.000*** (0.734)	[0.000, 6.086] (0.000, 7.349)	[0.507, 2.941] (0.043, 3.621)
College Grad vs HS Dropout	37.32 (12.54)	7.768*** (1.615)	[0.000, 7.777] (0.000, 10.807)	[0.581, 8.856] (0.115, 12.280)

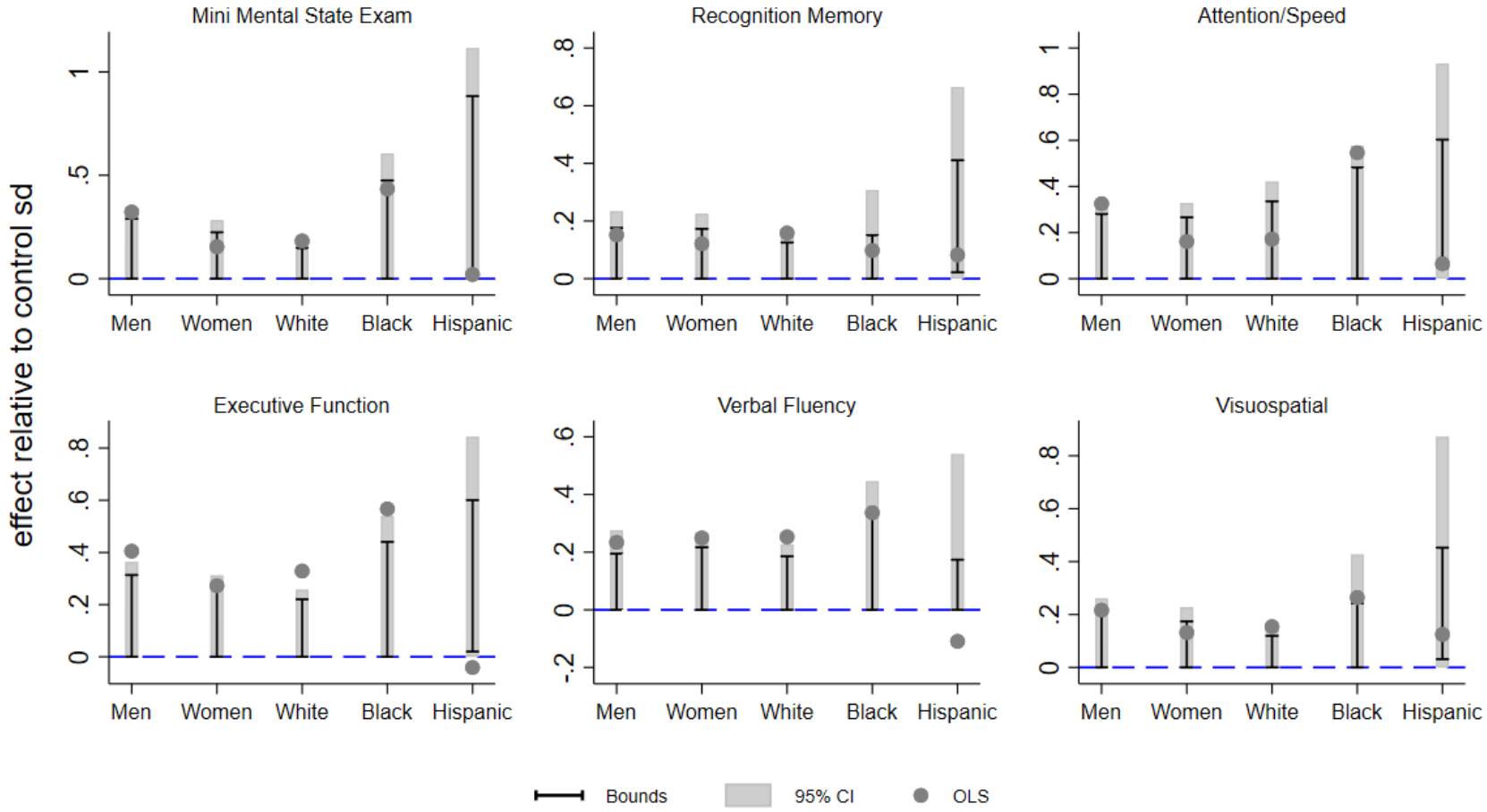
Notes: Robust standard errors in (.) in column 2. In columns 3 and 4 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. The min and max values of the residuals were used in computing the bounds. The MIV is mother's schooling with bins for high-school dropout, high-school graduate and more than high school. Analysis limited to individuals aged 25-50 years in the MIDUS Cognition Project. Estimation sample size in panels A, B and C are 1347, 1400, and 1393 observations respectively.

Figure C1: MTS+MTR+MIV Bounds on Effect of Completing High School on Cognition By Gender and Race



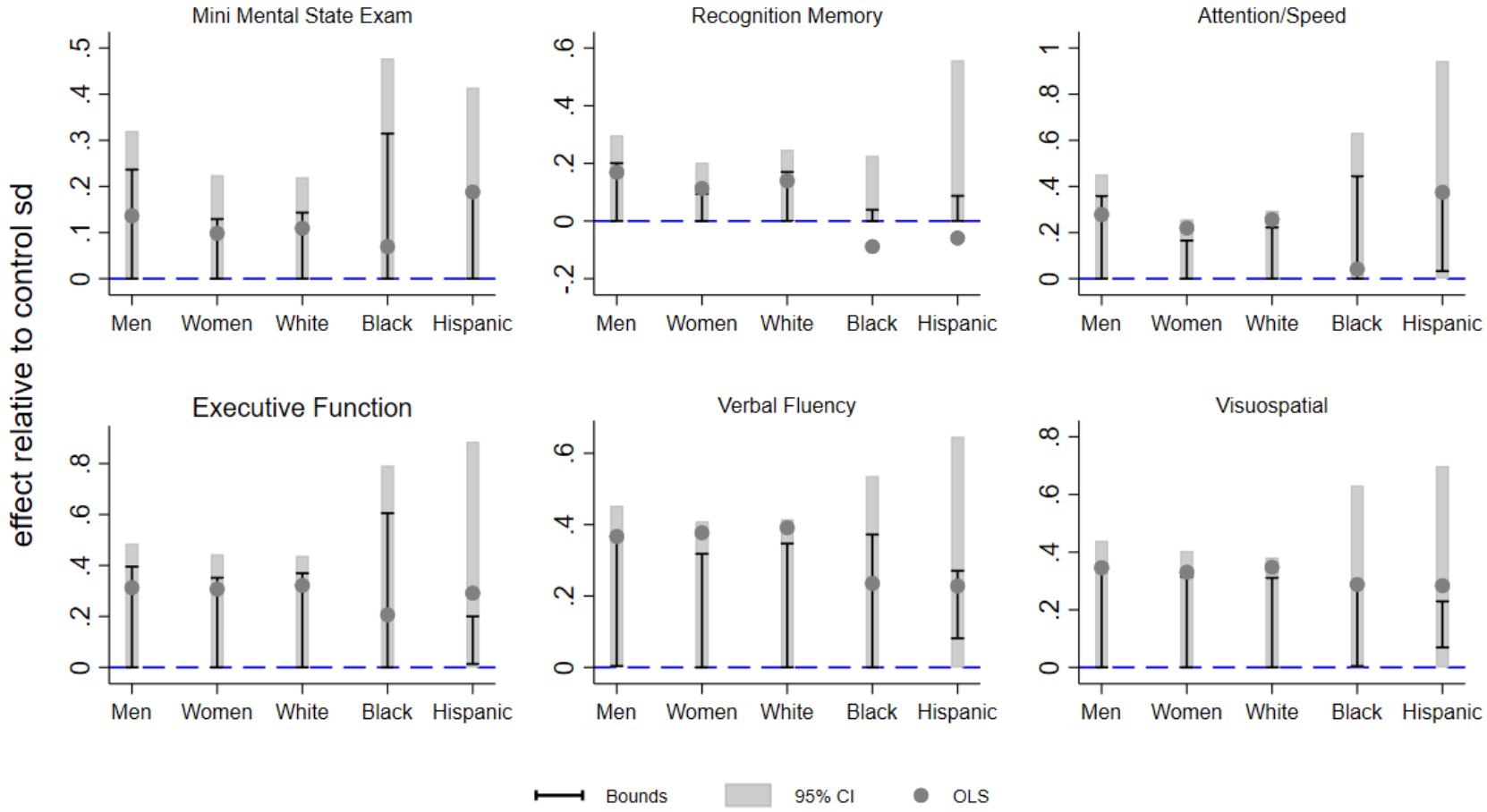
Notes: The figure presents the estimated bounds on the effect of going from being a high-school dropout to high-school graduate on cognition, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control (high-school dropouts)" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A.

Figure C2: MTS+MTR+MIV Bounds on Effect of Increasing Schooling from High School to Some College on Cognition By Gender and Race



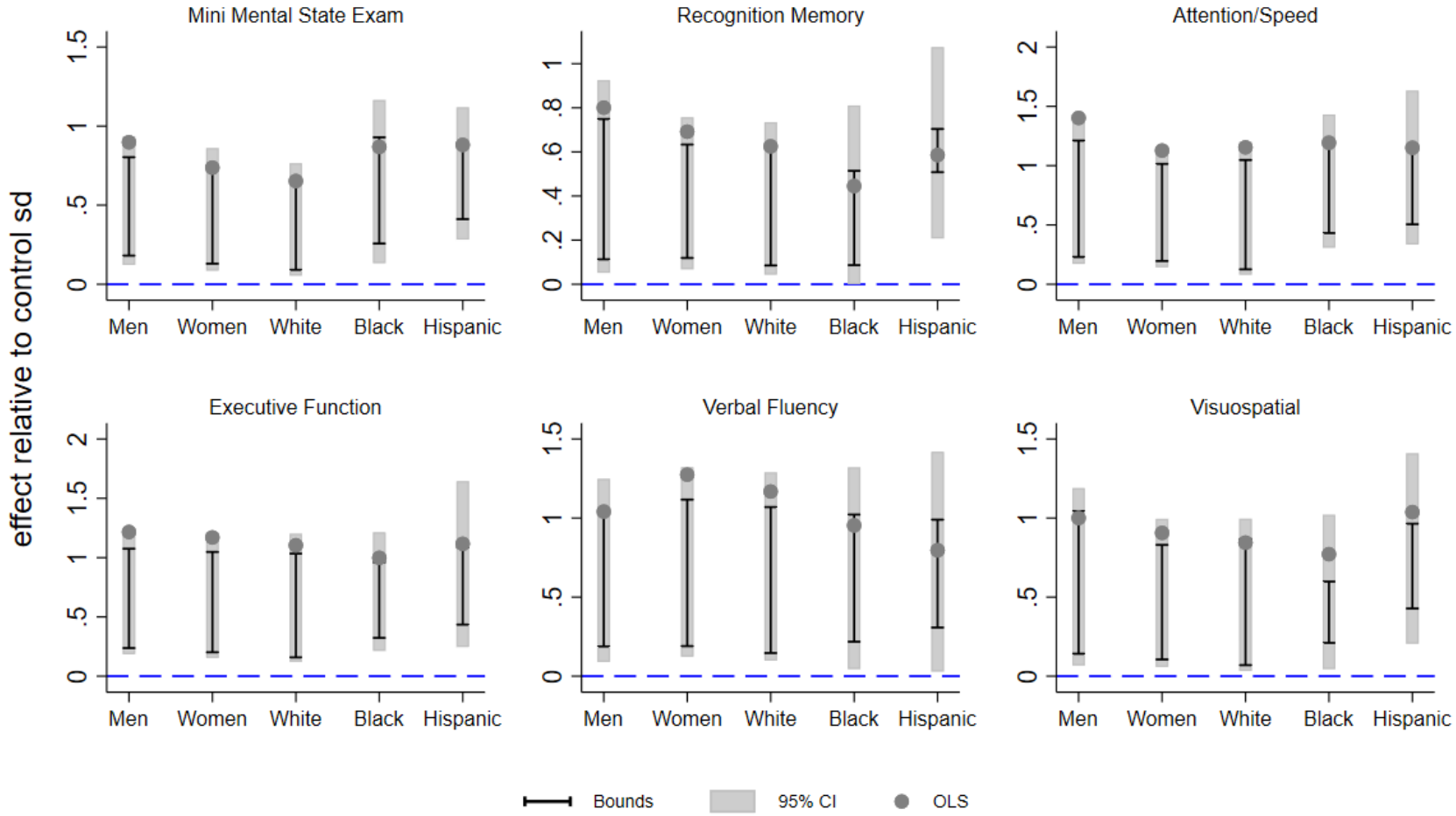
Notes: The figure presents the estimated bounds on the effect of going from being a high school graduate to having some college education on cognition, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control (high-school graduates)" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A.

Figure C3: MTS+MTR+MIV Bounds on Effect of Increasing Schooling from Some College to College Graduation on Cognition By Gender and Race



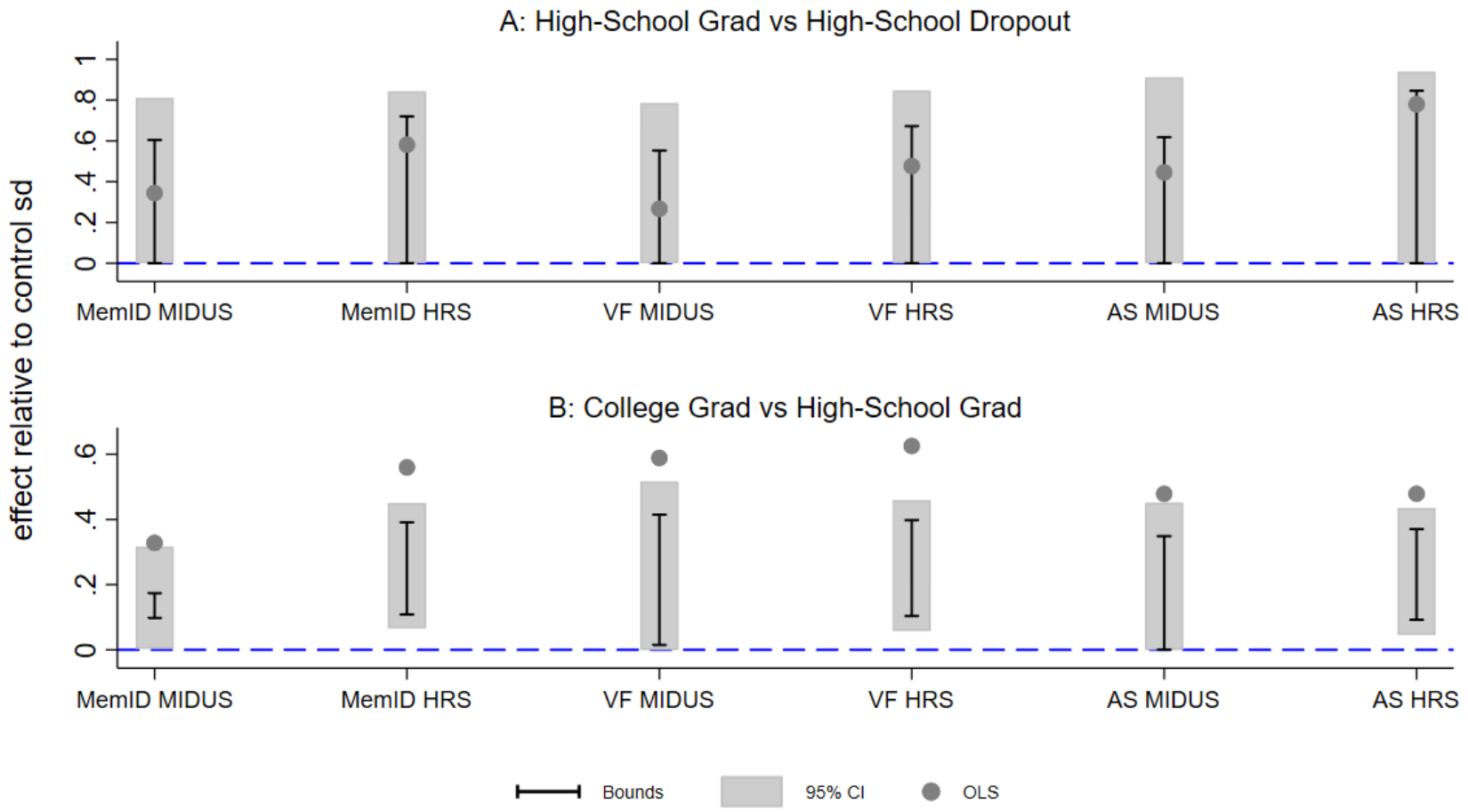
Notes: The figure presents the estimated bounds on the effect of going from having some college education to being a college graduate on cognition, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control (some college)" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A.

Figure C4: MTS+MTR+MIV Bounds on Effect of Increasing Schooling from High-School Dropout to College Graduation on Cognition By Gender and Race



Notes: The figure presents the estimated bounds on the effect of going from being a high-school dropout to a college graduate on cognition, by gender and race. The y-axis indicates the size of the effects relative to the corresponding "control (high-school dropouts)" sd. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A.

Figure C5: Comparison of MTS+MTR+MIV Bounds in the HRS and MIDUS



Notes: The figure presents the estimated bounds on the effects of completing secondary schooling in panel A and increasing schooling from secondary to tertiary schoolin in panel B. The y-axis indicates the size of the effects relative to the corresponding "control" sd. The control group refers to high-school dropouts in panel A and high-school graduates in panel B. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the Chernozhukov et al. (2013) method described in appendix A. MemID; Immeddiate+Delayed Memory VF:Verbal Fluency; AS:Attention/Speed