Published in final edited form as:

Soc Sci Med. 2015 February ; 127: 159–170. doi:10.1016/j.socscimed.2014.06.049.

Can Education Rescu Genetic Liability for Cognitive Decline?¹

C. Justin Cook and

University of Wisconsin-Madison, 1180 Observatory Drive, Madison WI 53706

Jason M. Fletcher

University of Wisconsin-Madison, 1180 Observatory Drive, Madison WI 53706

C. Justin Cook: cjcook@ssc.wisc.edu; Jason M. Fletcher: jmfletcher@wisc.edu

Abstract

Although there is a vast literature linking education and later health outcomes, the mechanisms underlying these associations are relatively unknown. In the spirit of some medical literature that leverages developmental abnormalities to understand mechanisms of normative functioning, we explore the ability of higher educational attainments to “rescue” biological/genetic liabilities in brain function through inheritance of a variant of the *APOE* gene shown to lead to cognitive decline, dementia, and Alzheimer's disease in old age. Deploying a between-sibling design that allows quasi-experimental variation in genotype and educational attainment within a standard gene-environment interaction framework, we show evidence that the genetic effects of the “risky” *APOE* variant on old-age cognitive decline are absent in individuals who complete college (vs. high school graduates). Auxiliary analyses suggest that the likely mechanisms of education are most consistent through changing brain processes (i.e. “how we think”) and potentially building cognitive reserves, rather than alleviating old age cognitive decline through the channels of higher socioeconomic status and resources over the life course.

1. Introduction

The impacts of educational attainments on a variety of outcomes over the life course are large and well known. In addition to large increases in material resources (e.g. lifetime income) attributable to higher educational attainment, health status has been shown to be highly associated with education across time periods, across countries, and over the life

¹The authors also acknowledge co-funding from the National Institute of Child Health and Human Development and the Office of Behavioral and Social Sciences Research (OBSSR) (1R21HD071884).

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 AG-21079 and AG-033285), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. Since 1992, data have been collected by the University of Wisconsin Survey Center. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at <http://www.ssc.wisc.edu/wlsresearch/data/>. The opinions expressed herein are those of the authors.

© 2014 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

cycle. More highly educated mothers give birth to healthier babies (Currie and Moretti 2003) and more highly educated individuals live longer than individuals with lower levels of schooling; for example, the age-adjusted mortality rate of high school dropouts ages 25 to 64 was more than twice as large as the mortality rate of those with some college (table 26, National Vital Statistics Reports, 2001, Cutler and Lleras-Muney 2006). There is a large literature using changes in compulsory schooling laws in the 1900s to examine impacts of educational attainment on old age mortality. This literature has been quite mixed, with Lleras-Muney (2005) showing some evidence of effects in a US sample, but other studies in European countries showing no impacts. See Fletcher (2013) for a review and new evidence. In between birth and death, more highly educated individuals smoke less (Farrel and Fuchs 1982, Maralani 2013), are less likely to be overweight (McClaren 2007, Cutler and Lleras-Muney 2010), and are more likely to pursue preventative health care steps (Fletcher and Frisvold 2009). However, as methods aimed at causal inference have been employed, the evidence linking educational attainment and health status and behaviors has become more mixed (Royer and Clark 2013)

While there are large literatures examining the impacts of education on health behaviors and health status over time and across countries, the mechanisms underlying these links remain unclear. Indeed, a next step in understanding long term impacts of education on health is in considering specific mechanisms. One dichotomy that might help us understand the extent of key mechanisms is between socioeconomic and biological channels. Education may enhance future health through the acquisition of financial and social resources that are important for maintaining health (e.g. income, health insurance, strong peer networks) and/or it may enhance future health through structuring and re-structuring brain development and activity that is helpful for health and wellbeing (see Culter and Lleras-Muney 2006 for a review). Both are likely important channels, but the latter has had limited examination, particularly in explorations that have strong causal grounding.

This paper focuses attention on the second, biological, channel while attempting to hold the other channel constant in the context of a specific marker of health: life course brain function atrophy (i.e. cognitive decline). In order to uncover novel evidence of potential mechanisms underlying the relationship of education and cognitive function in old age, we focus attention on the well-known differences in trajectories of brain malfunction between individuals with alternative variants of the *APOE* gene. In particular, we ask whether higher educational attainment “rescues” genetic liabilities of cognitive decline in old age by enhancing cognitive reserve. To explore this question, we use unique panel data collected over 50 years, a gene-environment interaction framework and a sibling-difference specification. In doing so, we attempt to go “under the scalp” in examining mechanisms of educational attainment impacts. Indeed, we find evidence that, among college graduates, *APOE* differences do not lead to cognitive decline differences; among high school graduates, *APOE* differences lead to large differences in cognitive decline in old age. These findings do not change when we add potential “social” (i.e. non-biological) mediators, such as wealth, marital status, health insurance, occupation, etc., which is consistent with a biological mechanism linking education with cognitive reserve through changes in “how we think”.

2. Background

2.1 APOE4

The *APOE* gene is associated with the production of apolipoprotein, which transports cholesterol and other fatty acids within the blood (Bu 2009). The functional variation in *APOE* is the result of two SNPs, or singular nucleotide polymorphisms: SNP rs429358 and SNP rs7412, with each SNP having two alleles, or genetic variants. Three major functional variants exist for the *APOE* gene: APOE2, APOE3, and APOE4. For European populations, the respective allele distribution is roughly 14%, 72%, and 14% for the three variants (Singh et al. 2006).

The E4 variant, the variant of interest throughout the paper, is strongly associated with late-onset Alzheimer's Disease (LOAD), which occurs between 60 and 70 years of age (Blacker et al. 1997). For meta and genome wide association analyses of the association between LOAD and APOE4 see Saunders et al. (1993), Farrer et al. (1997), and Bertram et al. (2007). While roughly 15% of the general population possesses the E4 variant, the frequency rises to roughly 40% in those with Alzheimer's Disease (Corder et al. 1993).

One potential mechanism of APOE's role in LOAD is in the accumulation of amyloid plaques (Bu 2009). Amyloid precursor proteins are hypothesized to play a role in synapse formation, and the accumulation of a byproduct of this protein, beta amyloid, has strong associations with AD (Blennow et al. 2006, Priller et al. 2006). Compared to the more common E3 variant, the E4 variant of APOE is less efficient at removing beta amyloid, leading to a greater accumulation of harmful amyloid plaques (Bu 2009). A number of mouse studies confirm the poor clearance of E4 for the beta amyloid peptide; see e.g., Holtzman et al. (1999), Holtzman et al. (2000), and DeMattos et al. (2004).

The timing of the impacts of the E4 variant is important. Because the less efficient polymorphism allows the greater accumulation of plaques over the life course, the impacts of having the “risk” allele are not apparent until old age. Specifically, this means that educational attainments and cognitive function during adolescence and young adulthood are likely not to be impacted. Like many other studies, we show this in our data—individuals with the E4 variant have the same IQ at age 17 and have the same educational attainments as individuals with an alternative variant. This is consistent with evidence from Ilhe et al. (2012), from which the authors find no association between the harmful E4 variant and early-life cognitive function. The accumulation of amyloid plaques, which is associated with later-life loss of cognitive function, occurs throughout the life-course and materializes in the late-onset period of 60-70 years of age. The accumulation of beta amyloid is hypothesized to affect cognition 2-3 decades prior to the onset of AD, a time after the formal education period (Davies et al. 1988, Villemagne et al. 2013). This particular timing of effects of the E4 variant over the life course can allow a unique lens in understanding the role of education in cognitive function and decline, as well as assessing causality that have not been exploited for these purposes in the literature. In order to pursue these questions, we take advantage of the emerging gene-environment interaction framework.

2.2 Gene-Environment Interaction

A growing literature is focused on the differential response to environmental stimuli based on underlying genetic differences within individuals. These interactions between genes and environment provide evidence for the moderating, or amplifying, influence of certain genetic variants in explaining heterogeneity in health, cognitive, and economic outcomes from exposure to harmful or beneficial environments (for review see Caspi and Moffit 2006). An alternative view of this research is to focus on the moderating influence of environmental exposures to a harmful genetic variant, which are strongly associated with an observed, or phenotypic, outcome. In other words, the negative outcomes, which are the result of genetic endowments determined at conception, can be reversed by exposure to particular environments. With this idea in mind, we focus on the role of *APOE4* in explaining declines in later-life cognition.

As discussed above, Late-onset Alzheimer's Disease (AD), which typically occurs between 60 and 70 years of age, is strongly associated with the E4 variant of the apolipoprotein-E (*APOE*) gene (Rhinn et al. 2013). This association is one of the most widely recognized and replicated instances of a singular genetic change being associated with an observed behavior, or phenotype (see e.g., Bertram et al. 2007 for meta-analysis and the resulting AlzGene database). Individuals with two copies of the E4 variant have been shown to be 7 times more likely to develop AD than those with the more common E3 variant (Corder et al. 1993). The association between *APOE4* and cognition does not exist, however, early in life, suggesting that any beneficial environmental experiences are unlikely to be driven by genetic variation in *APOE* (Ihle et al. 2012). This is important from a research design perspective, as gene-environment correlation (“genes selecting environments”) can challenge attempts at estimating causal impacts of gene-environment interactions (Fletcher and Conley 2013).

Towards this end, we propose that formal education serves as a moderating factor in the expression of the E4 variant for later-life declines in cognition. Physiologically, years of schooling has been shown to increase the volume and metabolism of grey matter while also strengthening neurological connections (Arenaza-Urquijo et al. 2013). Additionally, cognitive stimulation in early to mid-life (a time span correlated with the formal education period) has been shown to reduce the accumulation of amyloid-beta deposition in later-life (Landau et al. 2012).

Our proposed hypothesis is that the negative effects of *APOE4* on later-life cognition are offset by increases in education, measured by years of schooling. Years of schooling represents an environmental “shock” (i.e. unrelated to genotype) in early life that has effects on both the physiological development of the brain and in unobserved cognitive processing. Towards this end, we estimate a gene-environment interaction model between the harmful, or cognitively damaging, variant of the *APOE* gene and years of schooling on changes in later-life cognition during the late-onset period of AD. Given this estimation strategy our focus is on the marginal effect of *APOE4* for varied levels of schooling, with the hypothesized effect being a lessened impact of the harmful E4 variant for individuals with increased levels of schooling. Furthermore to lessen potential bias from unobserved

environments as well as the unobserved portion of the genome, our estimation strategy employs sibling fixed effects. The importance of this strategy is highlighted in Cook and Fletcher (2013), Conley and Bennet (2000), Conley et al. (2003), and Conley and Rauscher (2013). The use of sibling fixed effects further allows the interpretation of leveraging a random assignment (i.e. “genetic lottery”) (Fletcher and Lehrer 2009, 2011) of the E4 endowment at conception.

There is some related research on this question in the epidemiological literature, though the findings have been quite mixed. Using a sample of elderly individuals, Shadlen et al. (2005) show that having two copies of the E4 variant is more strongly related to later-life cognition of those with lower levels of formal education. Shalden et al. (2005) find no significant interaction for individuals possessing only one copy of the E4 allele. An additional study, using data from the MacArthur Successful Aging Study, provides evidence against our proposed hypothesis as well as that in Shadlen et al. (2005). Estimates from Seeman et al. (2005) provide evidence that individuals with greater than 8 formal years of education experience greater losses in cognition over time, implying education is serving as an amplifier for the harmful effects of the E4 variant. A further study, however, finds mostly insignificant associations of the interaction between APOE4 and education on declining cognition in later life (Van Gerven et al. 2012). Whether the differences in findings reflect differences in study samples (country, ages) or empirical methodologies, is not clear. Our approach pushes this literature forward by focusing on a sibling-difference empirical design, examining alternative mechanisms, and leveraging over 50 years of data.

We are able to show that the moderating influence of education is not the byproduct of increased incomes, better access to medical care, job characteristics, and other personal or social factors related both to education and cognitive decline. This provides further support for our main hypothesis that formal education has a direct effect on cognitive functioning which serves as a hedge to the harmful effects of the E4 variant of the APOE gene.

2.3 Cognitive Reserve

Our hypothesis that years of schooling lessen the effects of *APOE4* is tied closely to the idea of cognitive reserve. The hypothesis of cognitive reserve implies a stock of cognition that can be degraded and adjusted from both the environment and the genome (Stern 2009, 2012). In other words, heterogeneity exists across individuals in the ability to maintain normal cognitive function from an equivalent shock.

Cognitive reserve has two theoretical components. The first alludes to an unchangeable endowment imbued in early life; this is referred to as brain reserve (Katzman 1993, Stern 2012). The other component of cognitive reserve, which is simply referred to as cognitive reserve, represents the ability of the brain to adjust to environmental stimuli, such as education. A highly related idea is neuroplasticity (See Pascual-Leon et al. 2005 for review).

These two subcategories of cognitive reserve may be related. Individuals endowed with greater early-life cognitive abilities are likely to pursue cognitively rewarding outcomes like education. Therefore, in order to measure the transformative effects of education, it is necessary to control for the individual's “brain reserve endowment”.

3. Data and Empirical Methodology

3.1 Data

The data to be used are from the Wisconsin Longitudinal Study (WLS), which is comprised of a random sample of one-third of the 1957 high school graduating class within Wisconsin. Subsequent waves also collected data on a singular selected sibling of the originally sampled graduate. The WLS is a life-long panel, for which genetic data were collected in the 2003 wave (2004 wave for siblings). In addition to the genetic data and a wide array of individual-level data, the WLS began testing individuals for cognition in the 1992/1993 wave.

Our primary outcome variable of interest is an indicator for having either no change or a positive change in cognition between the 2003/2004 wave and the 2011 wave. This time frame corresponds to the onset of LOAD and is an ideal time to measure changes in cognition due to the E4 genetic variant (Rhinn et al. 2013). In measuring cognition, we use identical cognition test scores found within the 2003/2004 wave and the 2011 wave. These measures of cognition are based on letter fluency, similarities, and word recall. Scores for the three tests are aggregated to form a cognition score for both the 2003/2004 wave and the 2011 wave. The scores contain differing numbers of questions. In combining the three cognition scores into a single test score, we take the average of the percent of correct answers for each test. Next, the percentage change between the two considered periods is calculated, and an indicator variable is created for those who did not experience a decline in cognition over this period. Given the limited available data on cognitive differences between the two waves of interest, the three measures of cognition may not truly account for meaningful cognitive changes; therefore, the estimation of our gene-environment interaction may not be conclusive. Additionally, we regress each individual measure that makes up our base cognition score on the gene-environment interaction of interest in Tables A6-A8.

Our hypothesis focuses on the moderating influence of education in explaining the harmful effects of the E4 variant. To measure education, we use self-reported years of schooling, which is standardized within the sample for ease in interpreting regression coefficients. Data for years of education comes from the 1992/3 wave, a time later in life from which additional years are unlikely to be accumulated. The E4 allele is determined by one of four genetic variants for the APOE gene. These four variants are the product of two SNPs, or singular nucleotide polymorphisms—a singular change in a strand of DNA. The two SNPs are rs429358 and rs7412, from which the E4 variant is defined as a “C” variant for each SNP. In measuring the presence of E4 variants, we use the count of the number of each E4 variant within an individual's genotype. Alternative measures—i.e., an indicator for possessing at least one E4 variant—produce similar results (see appendix tables).

Genetic data for the WLS were collected in the 2003/2004 wave. For the two SNPs of APOE, data exists for roughly 4,400 graduates and 2,400 selected siblings. This sample, however, is truncated due to the availability of cognition scores for the 2003/2004 wave and the 2011 wave as well as data for years of schooling, which are needed to perform our hypothesized estimation. After this truncation, our sample contains data for roughly 3,400 individuals—2,400 graduates and 1,000 selected siblings. However, in order to perform our baseline estimation, which includes family-level fixed effects, we need complete data on

sibling pairs, which further truncates our sample. Our base sample contains complete data for 934 individuals, or 467 sibling pairs.

The truncations of our base sample may not be random. DNA data, which were collected in the 2003/2004 wave, occur at a time in which most individuals in the WLS are in their mid-sixties, implying survival into the initial sample may not be random and may be correlated with either education or cognition. This problem of sample selection is likely further exacerbated by the further reduction to complete sibling pairs. To account for this issue, attrition weights are calculated based on IQ as well as other demographic factors (essentially our demographic controls).

An additional issue with the WLS is in the homogeneity in ethnicity of the sample: our base sibling sample consists only of individuals of European decent. Therefore, generalizations of our findings to more ethnically diverse populations as that in the U.S. should be tempered; however, the lack of ethnic diversity within our base sample does alleviate concerns associated with genetic and cultural clustering within ethnicity (i.e. population stratification).

3.2 Empirical Methodology

Our primary estimating equation is given by the following form:

$$Cog_{ij} = \beta_0 + \beta_1 APOE_{4ij} + \beta_2 EDU_{ij} + \beta_3 APOE_{4ij} \times EDU_{ij} + \beta_4 X_{ij} + \beta_5 Z_{ij} + \gamma I_j + \varepsilon_{ij} \quad (1)$$

where we consider i individuals within j families. Our main outcome of interest is an indicator for not experiencing a decline in cognition, Cog_{ij} , and the coefficient of interest is β_3 , which measures the effect on the interaction between the number of E4 variants of the APOE gene and a standardized measure for the years of formal education an individual has. Our hypothesis being that β_3 is positive and significant, while the main effect of APOE4, measured by β_1 , is negative and significant. This finding would confirm that the effects of APOE4 on later life cognition are moderated by increasing levels of education. All estimations also include individual demographic controls—represented by X_{ij} —that include the initial cognition score from the 2003/2004 wave, high school IQ—a proxy for cognitive endowment, birth year, an indicator for sex, and birth order, which is shown to have effects both on early-life learning and cognition (Black et al. 2005).

The inclusion of family-level, or sibling, fixed effects is represented by γ_j . The fixed effects specification represents our base model. Controlling for unobserved family level factors provides two benefits. First, the inclusion of sibling fixed effects allows us to control for unobserved environments shared between siblings—e.g., habits, values, diet, etc.—that may be correlated with either educational attainment or later-life changes in cognition. The second benefit from the inclusion of sibling fixed effects is due to the fact that siblings share roughly 50% of unique genetic variation. Therefore, the inclusion of sibling fixed effects is able to account for large, unobserved portions of the genome, which can potentially drive either the level of schooling an individual obtains or other traits associated with cognition. Additionally, within-family estimation randomizes the genetic treatment, where each sibling

has equal odds of obtaining a particular genetic variant (i.e. the “genetic lottery” as discussed in Fletcher and Lehrer 2009, 2011).

Although we are able to potentially reduce bias from unobserved, time invariant family environments, we are not able to account for individual specific effects that may be associated with both educational attainment and our outcomes of interest. To address this issue, our set of baseline controls, particularly IQ, attempts to account for these individual-level differences, from which roughly 22% of the variation in years of schooling between siblings is accounted for from the set of baseline controls, implying large amounts of the variation in education across siblings is unobserved. Additionally, genetic differences across siblings remain. As stated above, roughly 50% of this variation is accounted for, but the remaining variation may have associations with educational attainment. Indeed in a study of siblings, Rietveld et al. (2013) find a number of within-family genetic associations with years of schooling, though the amount of variation in education explained by the significant genetic variants is approximately 2%. We consider the alternative hypothesis of unmeasured gene-gene interaction as the primary explanation of our findings to be unlikely (indeed, we know of no evidence of gene-gene interactions of any sizable magnitude that have been found in the literature), but we cannot rule this alternative hypothesis out until genome-wide data is available for the WLS sample.

Estimated coefficients of the proposed regression specification are given in the next section. All tables follow the form: column (1) performs estimation with the large as possible sample, column (2) repeats the estimation of column (1), restricting the sample to our base sibling sample; column (3) re-estimates column (2), adjusting for the inverse probability of attrition; and column (4) estimates a fixed-effects model.

4. Results

4.1 Baseline Results

Table 1 gives the main effects of both years of schooling (adjusted to a standard normal distribution) and APOE4 in measuring the probability of not experiencing a decline in cognition between the 2003 (2004 for sibling) and 2011 waves of the WLS. As is shown in the coefficient, education has a positive and statistically significant effect on the probability of not experiencing a decline in cognition and this effect is roughly consistent throughout the empirical specifications of Table 1, all of which include our baseline set of controls. From our baseline estimation of column (4), which performs within-family estimation, an increase in the years of schooling by one standard deviation (roughly 2 years) is associated with an increase in the probability of having constant or improving cognition of 8 percentage points. The E4 variant also has a statistically significant and consistent magnitude from estimation in Table 1; however unlike schooling, the E4 variant has a negative association with cognition. The magnitude of the coefficient of the count of E4 variants is significantly larger for within-family estimation. From the within-family estimation of column (4), possessing one copy of the E4 allele is associated with a decline in the probability of having constant or improving cognition by 14 percentage points. As expected, education has a positive association with cognitive outcomes while the E4 variant has a negative association.

Our focus, however, is not on main effect of each variable in explaining cognition; rather, our main hypothesis is that the negative effects of APOE4 on cognition are moderated by increased education. Before exploring how the effect of APOE4 is moderated by education, we must first ensure that years of schooling and early life cognition are not driven by APOE4. Table 2 regresses both years of schooling (Panel A) and early life cognition (measured by high school IQ; Panel B) on APOE4. For both outcomes, APOE4 has a statistically insignificant effect, implying that our genetic endowment of interest is not a potential source of the variation in our environment of interest. This result is to be expected as APOE4's hypothesized role in cognitive decline is due to an accumulation of amyloid plaques over the life course. Early life events, particularly human capital accumulation, are unlikely to be influenced by the E4 variant (Ilhe et al. 2012).

4.2 Gene-Environment Interaction

Our baseline estimating equation, outlined in Section 3.2, is estimated in Table 3. Our focus is on the marginal effect of APOE4 for differing levels of schooling, which is determined by the coefficient on both the main effect of the number of E4 variants and its interaction with years of schooling. The marginal effect for two levels of years of schooling is reported at the bottom of Table 3 as well as in subsequent tables. Marginal effects are reported for two levels of years of schooling: one standard deviation above and below the mean. Given that the mean is roughly 14 years and the standard deviation is roughly 2 years, these levels correspond to college graduates and high school graduates, respectively.

For column (1), which gives estimates for as large a sample as possible, all terms for the interaction model are statistically significant with the expected sign: the coefficient on our measure of APOE4 is negative, while the coefficients on years of schooling and the interaction term are positive. The positive coefficient of the interaction term implies that the negative effects APOE4 on cognition are lessened for more years of schooling. This is confirmed in the estimated marginal effect of APOE4. For individuals with high school or less, the number of E4 alleles is strongly and negatively associated with the probability of experiencing no or positive change in cognition. Interpreting the marginal effect in column (1), having one copy of the E4 variant reduces the probability of not having a decline in cognition by 8 percentage points, whereas having two copies reduces the probability by 16 percentage points. The magnitude of the marginal effect of APOE4, however, is reduced substantially for those who are college graduates, leading to a statistically insignificant association between APOE4 and our cognitive outcome of interest. The findings of column (1) support our main hypothesis.

Columns (2) and (3) provide simple OLS and weighted estimates, respectively, for our base sibling sample. The coefficient of the interaction in column (2) while remaining similar in magnitude to the estimate of column (1) loses statistical significance from a loss in precision in the smaller sibling sample. The marginal effect of APOE4, however, is consistent with the previous estimation: for those with 12 years of schooling or less, APOE4 has a strong negative association with cognition. This negative effect dissipates, however, for those with 16 or more years of schooling. The estimates of column (3), which weight estimation to

account for possible selection into our base sibling sample, are similar to those in column (2).

Finally, column (4) gives our base specification, which includes sibling fixed effects into the estimation of column (2). For the within-family estimation of column (4), the main effects of education and APOE4 as well as the interaction are all as expected with significant coefficients for both our measure of APOE4 and the interaction term. The marginal effect of APOE4 is similar to the previous estimates of columns (1)-(3), from which a negative and highly significant effect of APOE4 is seen for lower levels of education but statistical significance is lost when considering individuals with more years of schooling.

Table 4 re-estimates the findings of Table 4 while controlling for the distance from the mean of the change in cognition. Controlling for the distance from the mean is intended to correct for differences associated with relatively large versus small changes in cognition; while the mean change in cognition is negative for our base sibling sample, the distribution is slightly skewed to the right, implying increased variation in positive changes in cognition. Controlling for the magnitude of the change in cognition, however, does not alter the main findings of Table 3.

For all estimated coefficients in Table 4, magnitudes and significance are similar to equivalent estimation of Table 3. This is further seen in the marginal effect of APOE4 on the indicator for constant or improving cognition: for high school graduates, APOE4 has a negative and highly significant association with cognition. This effect, however, dissipates when considering college graduates.

Table 5 again replicates the estimation strategy of Table 3; although, in place of the indicator for not experiencing declining cognition, the percentage point change in cognition is considered as our dependent variable of interest. The use of the percentage point change is problematic, as we are not interested in the magnitude of the change in cognition, but rather the direction as an indication of cognitive impairment.

The estimates of Table 5 remain consistent in direction to those previously shown in Table 3; however, the coefficient of the interaction term is no longer statistically significant. For high school graduates and less, the marginal effect of APOE4 remains negative and highly significant for simple OLS for our base sibling sample. Interpreting the marginal effect of column (3), each additional copy of an E4 variant reduces an individual's change in cognition by roughly 4 percentage points if the individual has less than or equal to 12 years of education. This significant negative effect of APOE4, however, does not exist for individuals with 16 or more years of schooling.

For our base specification in column (4), which includes sibling fixed effects, point estimates remain roughly consistent in magnitude to the non-fixed-effects estimation of column (2)—this is seen in both the coefficients as well as the marginal effects of APOE4, but statistical significance dissipates. Given the consistence in magnitude of the coefficients, the loss in significance is attributed to the loss in variation from use of within-family estimation.

The estimated gene-environment interactions seen in Tables 3-5 support our main hypothesis. Greater levels of education, particularly levels of education corresponding to college graduation, lessen the effect of the harmful APOE4 variant in determining later-life cognition. Estimated marginal effects, as well as the coefficient of the main effect of APOE4, provide evidence that the statistically significant negative effect of the E4 in explaining later-life cognition dissipates as years of schooling increases above mean.

4.3 Potential Mechanisms

Given the previously shown moderating properties of education in explaining the relationship between APOE4 and later-life cognition, this section attempts to control for correlates of education which may be driving the interaction of interest. In order to parse the differential channels of education, we consider the interaction between a number of socioeconomic and behavioral outcomes associated with higher levels of education and APOE4. Findings are discussed within Table 6. All estimations within Table 6 use our base sibling sample while controlling for sibling fixed effects.

Column (1) replicates our base finding, given by column (4) of Table 3. Column (2) controls for both net worth of the graduate or sibling's family in the 2003 wave and a prestige score for the last or current job. Education is positively correlated with income. This greater level of income is likely to be positively associated with increased access and use of medical care. Therefore, it is plausible that the beneficial effect of education is being driven by increased income. In addition to income advantages, jobs that require more education are associated with more cognitively demanding activities, which may serve to limit cognitive decline. In controlling for both net worth and job prestige in column (2) we are attempting to control for this potential channel between education and cognitive decline. The inclusion of the additional covariates, however, does not substantially alter our estimated coefficients of interest or the marginal effect of APOE4 for varied levels of schooling.

Column (3) takes the same approach as column (2), replacing job characteristics with an index of access to medical care. Fletcher and Frisvold (2009) find that higher educational attainment increases preventive health behaviors in old age. Again, given the strong association between years of schooling and income, we would expect education to also be associated with better health coverage. Estimates including an index for self-reported access to medical care and its interaction with the number of E4 alleles are not substantially different from our baseline estimation given by column (1), implying access to health care is not the driving force of the moderating influence of education.

A large number of personal characteristics and their respective interaction with APOE4 are considered within column (4). These include indexes for the big five personality traits--openness, conscientiousness, extraversion, agreeableness, and neuroticism. Personality measures are from the 1992/3 wave, a time after schooling decisions have been made. It is therefore possible that the personality scores have been influenced by the previously obtained level of education. In addition to personality scores, column (4) also includes controls for the number of hours the graduate or sibling reads each week, an indicator of the individual's college plans at 16 years of age, the individual's body mass index, and indicator variables for smoking and drinking behaviors. Reading represents a personality trait that has

been shown to reduce cognitive aging and is correlated with years of schooling, whereas the inclusion of college plans is intended as a proxy for unobserved traits associated with desired, not actual, college attendance. Health behaviors are also likely correlated with educational attainment and may have effects on later-life cognition. As shown in column (4), however, differential personality traits are not the source of the rescuing effect of education in explaining cognitive decline from APOE4, as the coefficients of interest are similar to those found in the baseline estimation of Table 3.

Column (5) attempts to control for adverse social environments by considering an indicator for not being continuously married between the 2003/4 and 2011 waves. The inclusion of this dummy and its interaction with APOE4 does not substantially alter our base findings.

Column (6) considers two alternative measures for brain reserve: parents' education and high school IQ. Given that all estimations of Table 6 include sibling fixed effects, no main effect of parents' education can be estimated; however, the interaction with the randomly determined genetic endowment can be estimated. Although not reported in Table 6, the interaction between parents' education and either the graduate or sibling's APOE4 endowment is positive and statistically significant, implying that individuals from highly educated parents have a lessened harmful effect from the number of E4 variants. This ties into the idea of brain reserve, in that these individuals are likely to be endowed with greater cognitive abilities, and is a further cause for concern in that parental education may be capturing genetic endowments not shared between siblings. This unobserved genetic endowment is potentially associated with parental education, which is then shared or not amongst siblings, and may moderate the impact of the *APOE4* allele, implying the possibility of an unobserved gene-gene interaction as an alternative interpretation of our hypothesized gene-environment interaction. This potential gene-gene interaction is also seen in the significant negative effect of APOE4 for college graduates. Controlling for the interaction with parents' education, however, does not reduce the moderating impact of an individual's level of schooling: the coefficient on the interaction of interest remains positive and statistically indistinguishable in magnitude from the baseline estimate ($p=0.47$), although the coefficient of the interaction is no longer significant at conventional levels. Column (6) also includes the interaction between early-life IQ and the number of E4 variants. High school IQ is included within our baseline set of controls and is included within all columns of Table 6. The interaction with APOE4, however, is included within columns (6) and (7). Like the other measure of brain reserve, a positive interaction exists but not at the expense of years of schooling, which is hypothesized to measure acquired cognitive reserve. The piecemeal inclusion of either IQ and parents' education and the respective interactions with *APOE4* does not result in a loss of significance for the coefficient of the interaction between the number of E4 variants and years of schooling. Furthermore, the inclusion of both potential mechanisms along with all others of Table 6 does not lead to a loss in significance in the coefficient of interest. This is shown in column (7).

All potential mechanisms considered are included within column (7). The inclusion of all additional controls as well as their interaction with APOE4 does not alter our base finding: the number of E4 alleles has a strong negative, statistically significant effect on cognition for

those with high school or less. This effect, however, does not exist for those individuals with at least a college education. The moderating effect of education does not appear to be driven by measured income, personality, or early-life cognition traits, supporting our hypothesis of strengthened cognition directly from formal education.

5. Conclusion

This research examines the moderating influence of formal education in the relationship between the E4 variant of the APOE gene and later-life cognitive decline. We leverage the known biological processes underlying the life course patterns of cognitive decline for carriers of the variant in order to contribute a novel investigation of the possible causal mechanisms between education and later health outcomes. In the spirit of research from the medical sciences that focuses attention on developmental abnormalities to understand mechanisms of normative functioning, we extend this lens in our analysis of potential causal processes from increased educational attainments—to “get under the scalp”. Our specific hypothesis is that formal education has the potential to ameliorate the harmful effects of having the E4 variant of the APOE gene. Although there are at least two ways this process could unfold— through increased socioeconomic resources over the life course that could be used to reduce cognitive declines or through changes in the biological functioning of the brain itself—we view our results as most consistent with the latter channel.

To test this hypothesis, we focus on within-family estimation, which leads to random assignment of genetic variants. Furthermore, the use of sibling fixed effects allows us to control for unobserved environmental and genetic factors that may influence estimation. Towards this end, we are able to show that the harmful effect of the E4 variant in explaining declining later life cognition is statistically indistinguishable from zero for individuals with at least 16 years of schooling. In contrast, individuals with a high-school or less education have a statistically strong and negative association between the number of E4 variants and cognition.

As a further test of the proposed causative channel that education has direct effects on cognitive capacity, which hedge the harmful effects of the E4 variant, we examine a number of likely mechanisms. The inclusion of these additional controls and their interaction with APOE4 does not substantially alter the estimated relationship between years of schooling, the number of E4 variants, and an indicator for non-declining cognition. The estimates of Table 6 provide further support for the direct role of education in moderating cognitive decline from APOE4.

Appendix Tables

Table A1
Main Effects of *APOE4* and Years of Schooling for
Table 4

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.05***	0.06***	0.05***	0.08***
	(0.01)	(0.02)	(0.02)	(0.02)
Number of E4 Alleles	-0.04***	-0.06**	-0.06**	-0.14**
	(0.01)	(0.03)	(0.03)	(0.06)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Dist. from the Mean of Change in Cognition	Y	Y	Y	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.21	0.19	0.20	0.62

Table A2
Main Effects of *APOE4* and Years of Schooling for
Table 5

Dependent Variable: Percent Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Panel A: Main Effects				
Standardized Years of Schooling	0.03***	0.03***	0.03***	0.04***
	(0.00)	(0.01)	(0.01)	(0.01)
Number of E4 Alleles	-0.02***	-0.03**	-0.02**	-0.03
	(0.01)	(0.01)	(0.01)	(0.02)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.26	0.25	0.25	0.65
Panel A: Main Effects				

Table A3
Alternative Coding for *APOE4*: Indicator for Possessing
At Least One E4 Allele

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Panel A: Main Effects				
Standardized Years of Schooling	0.05***	0.06***	0.05***	0.07***
	(0.01)	(0.02)	(0.02)	(0.03)
Indicator for Possessing an E4 Allele	-0.04**	-0.07*	-0.07*	-0.17**
	(0.02)	(0.03)	(0.04)	(0.08)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.17	0.16	0.16	0.61
Panel B: Interaction				
Standardized Years of Schooling	0.04***	0.04**	0.03*	0.05
	(0.01)	(0.02)	(0.02)	(0.03)
Indicator for Possessing an E4 Allele	-0.05***	-0.07**	-0.08**	-0.18**
	(0.02)	(0.04)	(0.04)	(0.07)
G × E	0.03*	0.05	0.05	0.09*
	(0.02)	(0.03)	(0.03)	(0.05)
Observations	3421	934	934	934
R Sqr.	0.17	0.16	0.17	0.61
Marginal Effect of <i>APOE4</i> for High School Grads	-0.08***	-0.12**	-0.12**	-0.28***
(i.e., 1 s.d. below mean years of schooling)	(0.03)	(0.05)	(0.05)	(0.09)
Marginal Effect of <i>APOE4</i> for College Grads	-0.02	-0.03	-0.03	-0.09
(i.e., 1 s.d. above mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.09)

Table A4
Alternative Coding for Years of Schooling: Indicator
for Having 12 or Less Years of Sch

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Panel A: Main Effects				

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Indicator for High School Graduates and Less	-0.09*** (0.02)	-0.10*** (0.03)	-0.10*** (0.03)	-0.12** (0.05)
Number of E4 Alleles	-0.04*** (0.02)	-0.06** (0.03)	-0.06** (0.03)	-0.14** (0.06)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.16	0.16	0.17	0.61
Panel B: Interaction				
Indicator for High School Graduates and Less	-0.08*** (0.02)	-0.06* (0.04)	-0.06 (0.04)	-0.06 (0.06)
Number of E4 Alleles	-0.03 (0.02)	-0.01 (0.04)	-0.01 (0.04)	-0.05 (0.07)
G × E	-0.03 (0.03)	-0.12** (0.06)	-0.12** (0.06)	-0.21** (0.08)
Observations	3421	934	934	934
R Sqr.	0.16	0.16	0.17	0.61
Marginal Effect of <i>APOE4</i> for High School Grads (i.e., indicator for high school grad or less = 1)	-0.06*** (0.02)	-0.13*** (0.04)	-0.13*** (0.04)	-0.26*** (0.08)
Marginal Effect of <i>APOE4</i> for Some College and College Grads (i.e., indicator for high school grad or less = 1)	-0.03 (0.02)	-0.01 (0.04)	-0.01 (0.04)	-0.05 (0.07)

Table A5
Alternative Coding for both *APOE4* and Years of
Schooling

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Panel A: Main Effects				
Indicator for High School Graduates and Less	-0.09*** (0.02)	-0.10*** (0.03)	-0.10*** (0.03)	-0.12** (0.05)
Indicator for Possessing an E4 Allele	-0.04** (0.02)	-0.07* (0.03)	-0.07** (0.04)	-0.17** (0.08)

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.16	0.16	0.17	0.61
Panel B: Interaction				
Indicator for High School Graduates and Less	-0.08***	-0.06	-0.06	-0.06
	(0.02)	(0.04)	(0.04)	(0.06)
Indicator for Possessing an E4 Allele	-0.03	0.00	-0.00	-0.06
	(0.03)	(0.05)	(0.05)	(0.10)
G × E	-0.03	-0.15**	-0.14**	-0.23**
	(0.03)	(0.07)	(0.07)	(0.10)
Observations	3421	934	934	934
R Sqr.	0.16	0.16	0.17	0.61
Marginal Effect of <i>APOE4</i> for High School Grads	-0.06***	-0.15***	-0.14***	-0.29***
(i.e., indicator for high school grad or less = 1)	(0.02)	(0.05)	(0.05)	(0.08)
Marginal Effect of <i>APOE4</i> for Some College and College Grads	-0.03	0.00	-0.00	-0.06
(i.e., indicator for high school grad or less = 1)	(0.03)	(0.05)	(0.05)	(0.10)

Table A6
Baseline Estimation with Disambiguated Measure for Cognition

Dependent Variable: Indicator for Positive or No Change in Word Recall between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.03**	0.04**	0.04*	0.06*
	(0.01)	(0.02)	(0.02)	(0.03)
Number of E4 Alleles	-0.01	0.00	0.02	0.05
	(0.02)	(0.03)	(0.03)	(0.07)
G × E	0.02	0.02	0.01	0.09
	(0.02)	(0.03)	(0.03)	(0.05)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y

Dependent Variable: Indicator for Positive or No Change in Word Recall between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.04	0.04	0.04	0.51
Marginal Effect of <i>APOE4</i> for High School Grads	-0.03	-0.02	0.00	-0.04
(i.e., 1 s.d. below mean years of schooling)	(0.03)	(0.05)	(0.05)	(0.09)
Marginal Effect of <i>APOE4</i> for College Grads	0.02	0.03	0.03	0.13*
(i.e., 1 s.d. above mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.08)

Table A7
Baseline Estimation with Disambiguated Measure for Cognition

Dependent Variable: Indicator for Positive or No Change in Similarities between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.04***	0.04**	0.04*	0.06**
	(0.01)	(0.02)	(0.02)	(0.03)
Number of E4 Alleles	-0.00	-0.02	-0.01	-0.03
	(0.02)	(0.03)	(0.03)	(0.07)
G × E	0.01	0.01	0.02	0.03
	(0.02)	(0.03)	(0.03)	(0.05)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.07	0.08	0.07	0.58
Marginal Effect of <i>APOE4</i> for High School Grads	-0.01	-0.03	-0.03	-0.06
(i.e., 1 s.d. below mean years of schooling)	(0.03)	(0.05)	(0.05)	(0.08)
Marginal Effect of <i>APOE4</i> for College Grads	0.01	-0.01	0.00	0.00
(i.e., 1 s.d. above mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.08)

Table A8
Baseline Estimation with Disambiguated Measure for
Cognition

Dependent Variable: Indicator for Positive or No Change in Word Recall between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.01	0.02	0.02	0.02
	(0.01)	(0.02)	(0.02)	(0.02)
Number of E4 Alleles	-0.04***	-0.05**	-0.04*	-0.09*
	(0.01)	(0.03)	(0.02)	(0.05)
G × E	0.02*	0.02	0.02	0.02
	(0.01)	(0.02)	(0.02)	(0.03)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.05	0.07	0.08	0.60
Marginal Effect of <i>APOE4</i> for High School Grads (i.e., 1 s.d. below mean years of schooling)	-0.07*** (0.02)	-0.08* (0.04)	-0.06 (0.04)	-0.11 (0.07)
Marginal Effect of <i>APOE4</i> for College Grads (i.e., 1 s.d. above mean years of schooling)	-0.02 (0.02)	-0.03 (0.03)	-0.02 (0.03)	-0.07 (0.05)

References

- Arenaza-Urquijo E, Landeau B, La Joie R, et al. Relationships between years of education and gray matter volume, metabolism and functional connectivity in health elders. *Neuroimage*. 2013; 83:450–457. [PubMed: 23796547]
- Bertram L, McQueen M, Mullin K, et al. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nature Genetics*. 2007; 39(1):17–23. [PubMed: 17192785]
- Black, et al. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics*. 2005; 120(2):669–700.
- Blacker D, Haines J, Rodes L, et al. ApoE-4 and age at onset of Alzheimer's disease: The NIMH genetics initiative. *Neurology*. 1997; 48(1):139–147. [PubMed: 9008509]
- Blennow K, de Leon M, Zetterberg H. Alzheimer's disease. *The Lancet*. 2006; 368:387–403.
- Bu G. Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy. *Nature Reviews Neuroscience*. 2009; 10(5):333–344.
- Caspi A, Moffitt T. Gene-environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience*. 2006; 7(7):583–590.
- Conley D, Bennett. Is biology destiny? Birth weight and life chances. *American Sociological Review*. 2000; 65(3):458–467.

- Conley, Rauscher. Genetic Interactions with prenatal social environment: Effects on academic and behavioral outcomes. *Journal of Health and Social Behavior*. 2013; 54(1):109–127. [PubMed: 23475742]
- Conley, et al. *The Starting Gate: Birth Weight and Life Chances*. Berkeley, CA: University of California Press; 2003.
- Conley, et al. Twin differences in birth weight: The effects of genotype and prenatal environment on neonatal and post-neonatal mortality. *Economics and Human Biology*. 2006; 4(2):151–183. [PubMed: 16439189]
- Cook, Fletcher. Interactive effects of in utero nutrition and genetic inheritance on cognition: New evidence using sibling comparisons. *Economics and Human Biology*. forthcoming.
- Corder E, Saunders A, Strittmatter W, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993; 261(5123):921–923. [PubMed: 8346443]
- Currie J, Moretti E. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*. 2003; 118(4):1495–1532.
- Cutler D, Lleras-Muney A. Education and health: Evaluating theories and evidence. NBER Working paper 12352. 2006
- Cutler D, Lleras-Muney A. Understanding differences in health behavior by education. *Journal of Health Economics*. 2010; 29(1):1–28. [PubMed: 19963292]
- Davies L, Wolska B, Hilbich C, et al. A4 amyloid protein deposition and the diagnosis of Alzheimer's disease. *Neurology*. 1988; 38(11):1688–1693. [PubMed: 3054625]
- DeMattos R, Cirrito J, Parsadanian M, et al. ApoE and clusterin cooperatively suppress Abeta levels and deposition: Evidence that ApoE regulates extracellular Abeta metabolism in vivo. *Neuron*. 2004; 41(2):193–202. [PubMed: 14741101]
- Farrell P, Fuchs V. Schooling and health: The cigarette connection. *Journal of Health Economics*. 1982; 1(3):217–230. [PubMed: 10263956]
- Fletcher JM, Conley D. The Challenge of Causal Inference in Gene-Environment Interaction Research: Leveraging Research Designs from the Social Sciences. *American Journal of Public Health*. 2013 Aug 8; 103(8):e1–e4. Published online.
- Fletcher J, Frisvold D. Higher education and health investments: Does more schooling affect preventative health care use? *Journal of Human Capital*. 2009; 3(2):144–176. [PubMed: 22368727]
- Fletcher JM, Lehrer SF. Genetic lotteries within families. *Journal of Health Economics*. 2011; 30(4): 647–659. [PubMed: 21664708]
- Fletcher JM. New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. Working Paper. 2013
- Holtzman D, Bales K, Wu S, et al. Expression of human apolipoprotein E reduces amyloid-beta deposition in a mouse model of Alzheimer's disease. *The Journal of Clinical Investigation*. 1999; 103(6):R15–R21. [PubMed: 10079115]
- Holtzman D, Bales K, Tenkova T, et al. Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease. *PNAS*. 2000; 97(6):2892–2897. [PubMed: 10694577]
- Ilhe A, Bunce D, Kliegel M. APOE ϵ 4 and cognitive function in early life: A meta-analysis. *Neuropsychology*. 2012; 26(3):267–277. [PubMed: 22329498]
- Katzman R. Education and the prevalence of dementia and Alzheimer's disease. *Neurology*. 1993; 43:13–20. [PubMed: 8423876]
- Lleras-Muney A. The relationship between education and adult mortality in the United States. *The Review of Economic Studies*. 2005; 72(1):189–221.
- Maralani V. Educational inequalities in smoking: The role of initiation versus quitting. *Social Science and Medicine*. 2013; 84:129–137. [PubMed: 23466258]
- McLaren L. Socioeconomic status and obesity. *Epidemiologic Reviews*. 2007; 29:29–48. [PubMed: 17478442]
- Pascual-Leone, et al. The plastic human brain cortex. *Annual Review of Neuroscience*. 2005; 28:377–401.

- Priller C, Bauer T, Mitteregger G, et al. Synapse formation and function is modulated by the amyloid precursor protein. *The Journal of Neuroscience*. 2006; 26(27):7212–7221. [PubMed: 16822978]
- Rhinn H, Fujita R, Qiang L, et al. Integrative genomics identifies APOE ϵ 4 effectors in Alzheimer's disease. *Nature*. 2013; 500(7460):45–50. [PubMed: 23883936]
- Rietveld C, Medland S, Derringer J, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*. 2013; 340(6139):1467–1471. [PubMed: 23722424]
- Royer H, Clark D. The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*. 2013; 103(6):2087–2120.
- Seeman T, Huang M, Bretsky P, et al. Education and APOE- ϵ 4 in longitudinal cognitive decline: MacArthur studies of successful aging. *Journal of Gerontology*. 2005; 60B(2):74–83.
- Shadlen M, Larson E, Wang L, et al. Education modifies the effect of apolipoprotein epsilon 4 on cognitive decline. *Neurobiology of Aging*. 2005; 26(1):17–24. [PubMed: 15585342]
- Singh P, Singh M, Mastana S. APOE distribution in world populations with new data from India and the UK. *Annals of Human Biology*. 2006; 33(3):279–308. [PubMed: 17092867]
- Stern Y. Cognitive reserve. *Neuropsychologia*. 2009; 47:2015–2028. [PubMed: 19467352]
- Stern Y. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurology*. 2012; 11(11):1006–1012.
- Van Gerven P, Van Boxtel M, Ausems E, et al. Do apolipoprotein E genotype and educational attainment predict the rate of cognitive decline in normal aging? A 12-year follow-up of the Maastricht Aging Study. *Neuropsychology*. 2012; 26(4):459–472. [PubMed: 22642392]
- Villemagne V, Burnham S, Bourqueat P, et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *Lancet Neurology*. 2013; 12(4):357–367.

Table 1
Main Effects of *APOE4* and Years of Schooling on Cognition

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.05***	0.06***	0.05***	0.08***
	(0.01)	(0.02)	(0.02)	(0.03)
Number of E4 Alleles	-0.04***	-0.06**	-0.06**	-0.14**
	(0.02)	(0.03)	(0.03)	(0.06)
Controls				
Demographic	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.17	0.16	0.16	0.61

Notes: (i) The dependent variable is an indicator for having declining cognition between the 2003 and 2011 waves. The “Number of E4 Alleles” represents the count of E4 alleles—0, 1, or 2—an individual possesses. (ii) Demographic controls include the cognition score for the 2003 wave, a standardized value of early-life IQ, an indicator for sex, birth year, and birth order. (iii) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

Table 2
Relationship between IQ and APOE4

Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Panel A: Dependent Variable: Years of Sch.				
Number of E4 Alleles	-0.03	-0.06	-0.06	-0.02
	(0.03)	(0.07)	(0.07)	(0.12)
Sex Indicator	-0.33***	-0.20***	-0.21***	-0.25***
	(0.04)	(0.07)	(0.07)	(0.08)
Birth Year	0.02***	0.01*	0.01**	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Birth Order	-0.09***	-0.11***	-0.10***	0.04
	(0.01)	(0.02)	(0.02)	(0.06)
Sibling Fixed Effects	N	N	N	Y
Observations	3421	934	934	934
R Sqr.	0.05	0.03	0.03	0.68
Panel B: Dependent Variable: IQ				
Number of E4 Alleles	0.21	-0.27	-0.34	-0.06
	(0.51)	(0.91)	(0.91)	(1.52)
Sex Indicator	0.78	1.36	1.61*	-0.02
	(0.49)	(0.97)	(0.97)	(1.14)
Birth Year	0.54***	0.42***	0.42***	0.78***
	(0.08)	(0.14)	(0.14)	(0.23)
Birth Order	-1.05***	-0.89**	-0.83**	-1.49**
	(0.15)	(0.39)	(0.39)	(0.74)
Sibling Fixed Effects	N	N	N	Y
Observations	3421	934	934	934
R Sqr.	0.03	0.02	0.02	0.67

Notes: (i) The dependent variable for Panel A is years of schooling adjusted to a standard normal distribution. For Panel B, early-life IQ, adjusted to a standard normal distribution, is the dependent variable. The “Number of E4 Alleles” represents the count of E4 alleles—0, 1, or 2—an individual possesses. (ii) Demographic controls are included within the table. (iii) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

Table 3
Baseline Estimation: Interaction between *APOE4* and Years of Schooling

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.04***	0.04**	0.03*	0.05
	(0.01)	(0.02)	(0.02)	(0.03)
Number of E4 Alleles	-0.04***	-0.07**	-0.07**	-0.16***
	(0.02)	(0.03)	(0.03)	(0.06)
G × E	0.03**	0.04	0.05*	0.09**
	(0.01)	(0.03)	(0.03)	(0.04)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.17	0.16	0.17	0.61
Marginal Effect of <i>APOE4</i> for High School Grads	-0.08***	-0.11***	-0.12***	-0.25***
(i.e., 1 s.d. below mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.08)
Marginal Effect of <i>APOE4</i> for College Grads	-0.01	-0.03	-0.02	-0.07
(i.e., 1 s.d. above mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.07)

Notes: (i) The mean years of schooling is roughly 14 years and the standard deviation is roughly 2 years, implying that individuals a standard deviation above the mean are representative of college graduates while those one standard deviation below the mean are representative of high school graduates only. (ii) The dependent variable is an indicator for having declining cognition between the 2003 and 2011 waves. The “Number of E4 Alleles” represents the count of E4 alleles-0, 1, or 2-an individual possesses. “G × E” is the interaction between years of schooling, adjusted to a standard normal distribution, and the count of E4 alleles. (iii) Demographic controls include the cognition score for the 2003 wave, a standardized value of early-life IQ, an indicator for sex, birth year, and birth order (iv) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

Table 4
Baseline Estimation: Controlling for the Magnitude of Change in Cognition

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.04***	0.04**	0.03*	0.04
	(0.01)	(0.02)	(0.02)	(0.03)
Number of E4 Alleles	-0.05***	-0.07**	-0.07**	-0.16***
	(0.01)	(0.03)	(0.03)	(0.06)
G × E	0.03**	0.04	0.04	0.10**
	(0.01)	(0.03)	(0.03)	(0.04)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Dist. from the Mean of Change in Cognition	Y	Y	Y	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Squ.	0.21	0.19	0.20	0.63
Marginal Effect of <i>APOE4</i> for High School Grads	-0.08***	-0.11***	-0.12***	-0.26***
(i.e., 1 s.d. below mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.08)
Marginal Effect of <i>APOE4</i> for College Grads	-0.02	-0.03	-0.03	-0.06
(i.e., 1 s.d. above mean years of schooling)	(0.02)	(0.04)	(0.04)	(0.07)

Notes: (i) The mean years of schooling is roughly 14 years and the standard deviation is roughly 2 years, implying that individuals a standard deviation above the mean are representative of college graduates while those one standard deviation below the mean are representative of high school graduates only. (ii) The dependent variable is an indicator for having declining cognition between the 2003 and 2011 waves. The “Number of E4 Alleles” represents the count of E4 alleles-0, 1, or 2-an individual possesses. “G × E” is the interaction between years of schooling, adjusted to a standard normal distribution, and the count of E4 alleles. (iii) Demographic controls include the cognition score for the 2003 wave, a standardized value of early-life IQ, an indicator for sex, birth year, and birth order (iv) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

Table 5
Baseline Estimation: Effect of Interaction on Percent Change in Cognition

Dependent Variable: Percent Change in Cognition between 2003 and 2011				
Sample	All	Siblings		
	(1)	(2)	(3)	(4)
Standardized Years of Schooling	0.03***	0.03***	0.03***	0.04***
	(0.00)	(0.01)	(0.01)	(0.01)
Number of E4 Alleles	-0.02***	-0.03**	-0.03**	-0.03
	(0.01)	(0.01)	(0.01)	(0.02)
G × E	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Controls				
Demographic and Family SES	Y	Y	Y	Y
Sibling Fixed Effects	N	N	N	Y
Estimation				
Weighting by Prob. of Being in Sib Sample	N	N	Y	N
Observations	3421	934	934	934
R Sqr.	0.26	0.25	0.25	0.65
Marginal Effect of <i>APOE4</i> for High School Grads	-0.02***	-0.04**	-0.04**	-0.04
(i.e., 1 s.d. below mean years of schooling)	(0.01)	(0.02)	(0.02)	(0.03)
Marginal Effect of <i>APOE4</i> for College Grads	-0.01	-0.02	-0.02	-0.02
(i.e., 1 s.d. above mean years of schooling)	(0.01)	(0.01)	(0.01)	(0.02)

Notes: (i) The mean years of schooling is roughly 14 years and the standard deviation is roughly 2 years, implying that individuals a standard deviation above the mean are representative of college graduates while those one standard deviation below the mean are representative of high school graduates only. (ii) The dependent variable is the percentage change in cognition between the 2003 and 2011 waves. The “Number of E4 Alleles” represents the count of E4 alleles-0, 1, or 2-an individual possesses. “G × E” is the interaction between years of schooling, adjusted to a standard normal distribution, and the count of E4 alleles. (iii) Demographic controls include the cognition score for the 2003 wave, a standardized value of early-life IQ, an indicator for sex, birth year, and birth order (iv) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

Table 6

Potential Mechanisms

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011						
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized Years of Schooling	0.05 (0.03)	0.04 (0.03)	0.05 (0.03)	0.02 (0.03)	0.05 (0.03)	0.06* (0.03)
Number of E4 Alleles	-0.16*** (0.06)	-0.17*** (0.06)	-0.15** (0.06)	-0.17** (0.07)	-0.23*** (0.06)	-0.20*** (0.06)
G × E	0.09** (0.04)	0.10** (0.05)	0.10** (0.04)	0.10** (0.05)	0.10** (0.04)	0.06 (0.04)
Controls						
Demographic	Y	Y	Y	Y	Y	Y
Sibling Fixed Effects	Y	Y	Y	Y	Y	Y
Potential Mechanisms (× Number of E4 Alleles):*						
Job Characteristics						
Net Worth in 2011	N	Y	N	N	N	N
Prestige Score for Current/Last Job in 2003	N	Y	N	N	N	N
Access to Health Care						
Average Score for Access to Health Care Satisfaction	N	N	Y	N	N	N
Personality [†]						
Openness Index	N	N	N	Y	N	N
Conscientiousness Index	N	N	N	Y	N	N
Extroversion Index	N	N	N	Y	N	N
Agreeableness Index	N	N	N	Y	N	N
Neuroticism Index	N	N	N	Y	N	N
Reading (Hours per Week)	N	N	N	Y	N	N
Planned College Attendance at 16	N	N	N	Y	N	N
BMI (2003/4)	N	N	N	Y	N	N
Current Smoker (2003/4)	N	N	N	Y	N	N

Dependent Variable: Indicator for Positive or No Change in Cognition between 2003 and 2011							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Alcohol Symptom Count	N	N	N	Y	N	N	Y
(2003/4)							
Spouse							
Indicator for Being Unmarried between 2003 and 2011	N	N	N	N	Y	N	Y
Cognitive Endowment [‡]							
Average of Parents' Education	N	N	N	N	N	Y	Y
High School IQ	N	N	N	N	N	Y	Y
Observations	934	934	934	934	934	934	934
R Sqr.	0.61	0.61	0.61	0.64	0.62	0.62	0.65
Marginal Effect of <i>APOE4</i> for High School Grads	-0.25***	-0.28***	-0.25***	-0.28***	-0.27***	-0.26***	-0.32***
(i.e., 1 s.d. below mean years of schooling)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.08)
Marginal Effect of <i>APOE4</i> for College Grads	-0.07	-0.07	-0.06	-0.07	-0.08	-0.14*	-0.12
(i.e., 1 s.d. above mean years of schooling)	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)	(0.07)	(0.09)

Notes: (i) The mean years of schooling is roughly 14 years and the standard deviation is roughly 2 years, implying that individuals a standard deviation above the mean are representative of college graduates while those one standard deviation below the mean are representative of high school graduates only. (ii) The dependent variable is an indicator for having declining cognition between the 2003 and 2011 waves. The "Number of E4 Alleles" represents the count of E4 alleles-0, 1, or 2-an individual possesses. "G × E" is the interaction between years of schooling, adjusted to a standard normal distribution, and the count of E4 alleles. (iii) Demographic controls include the cognition score for the 2003 wave, a standardized value of early-life IQ, an indicator for sex, birth year, and birth order (iv) Standard errors are clustered at the family level with *, **, and *** representing significance at the 10, 5, and 1% significance level, respectively.

* The main effect and its interaction with the number of E4 alleles is included in each specified column.

[†] The mean value for BMI, smoking behavior, and drinking behavior is imputed for missing values of each variable. Indicator variables that account for missing values and their interaction with APOE4 are included in column (5) and (7).

[‡] Due to the shared values of parental education amongst siblings, only the interaction with APOE4 is included. High school IQ is included within our baseline set of controls and is included within all columns of Table 6; the interaction between IQ and APOE4 is included within columns (6) and (7).