I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10567

Genetic Ability, Wealth, and Financial Decision-Making

Daniel Barth
Nicholas W. Papageorge
Kevin Thom

FEBRUARY 2017

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10567

Genetic Ability, Wealth, and Financial Decision-Making

Daniel Barth
University of Southern California

Nicholas W. Papageorge
The Johns Hopkins University and IZA

Kevin Thom
New York University

FEBRUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

Genetic Ability, Wealth, and Financial Decision-Making*

Recent advances in behavioral genetics have enabled the discovery of genetic scores linked to a variety of economic outcomes, including education. We build on this progress to demonstrate that the same genetic variants that predict educational attainment independently predict household wealth in the Health and Retirement Study (HRS). This relationship is partly explained by higher earnings, but a substantial portion of this association cannot be explained mechanically by income flows or bequests. This leads us to explore the role of beliefs, financial literacy and portfolio decisions in explaining this genetic gradient in wealth. We show that individuals with lower genetic scores are more prone to reporting "extreme beliefs" (e.g., reporting that there is a 100% chance of a stock market decline in the near future) and they invest their savings accordingly (e.g., avoiding the stock market). Our findings suggest that genetic factors that promote human capital accumulation contribute to wealth disparities not only through education and higher earnings, but also through their impact on the ability to process information and make good financial decisions. The association between genetic ability and wealth is substantially lower among households receiving a defined benefit pension. Policies that transfer greater responsibility to individuals to manage their wealth might therefore exacerbate the consequences of labor market inequality.

JEL Classification: D14, D31, G11, H55, I24, J24

Keywords: wealth, inequality, portfolio decisions, beliefs, education and genetics

Corresponding author:
Nicholas W. Papageorge
Johns Hopkins University
Department of Economics
Wyman 521
3400 N. Charles St
Baltimore, MD 21218
USA

E-mail: papageorge@jhu.edu

* We thank Aysu Okbay for constructing and sharing the polygenic score used for HRS respondents. We thank Pietro Biroli for constructing and sharing the polygenic scores for cigarettes per day and BMI. We are also grateful for helpful comments from Jess Benhabib, Dan Benjamin, Alberto Bisin, Christopher Carroll, David Cesarini, Steven Durlauf, Jon Faust, Titus Galama, George-Levi Gayle, and Basit Zafar, along with seminar participants at the 2016 HCEO Conference on Genetics and Social Science. The usual caveats apply.

1 Introduction

Wealth inequality in the United States and many other countries is large and growing (Saez and Zucman, 2014; Jones, 2015). An extensive literature attempts to explain wealth inequality through variation in the earnings process, entrepreneurial talent, bequests, risk aversion and time-discounting. While these factors can generate a skewed wealth distribution, they fail to reproduce other key empirical features including the thickness of the tail.[1]

A new wave of theoretical work argues that cross-sectional heterogeneity in the returns to wealth is required to match the basic features of the wealth distribution (Benhabib, Bisin, and Zhu, 2011; Benhabib and Bisin, 2016). This argument is supported by a growing empirical literature that finds substantial heterogeneity in such returns (Fagereng et al., 2016; Benhabib, Bisin, and Luo, 2015; Bach, Thiemann, and Zucco, 2015). Much of this heterogeneity persists over time, with some individuals earning consistently higher returns to wealth (Fagereng et al., 2016). Despite the theoretical and empirical importance of this phenomenon, little is known about what drives such persistence. However, policy responses to wealth inequality are likely to have different effects depending on the mechanisms through which persistent heterogeneity operates — for example, whether this heterogeneity comes from variation in preferences (e.g. risk aversion) or variation in skills (e.g. financial decision-making)

In this paper, we identify a biological source of wealth inequality: genetic ability related to human capital accumulation. In particular, we show that the same observed genetic markers that predict educational attainment also predict household wealth in the Health and Retirement Study (HRS). Importantly, this relationship is not merely driven by earnings or other income flows. Rather, the estimated association between genes and wealth is economically large and statistically significant after conditioning on education, earnings, financial bequests and a host of other controls. Observing genetic variants implicated in wealth accumulation allows us to directly assess the mechanisms through which they operate and how they interact with policy-relevant environments. We provide novel evidence that the genetic endowments we study are linked to financial decision-making (in particular stock market participation), financial literacy, and probabilistic thinking.

Our results provide a genetic micro-foundation for persistent differences in returns to wealth needed to match existing wealth distributions. Biological heterogeneity may thus drive part of the association between wealth and the returns to wealth. Since information

[1] Some models can match the thickness of the tail, but only under implausible assumptions about the level of heterogeneity in the earnings process. For example, Kindermann and Krueger (2014) require the top 0.25% of earners to earn 400-600 times more than the median earner. Empirically, this number is closer to 33.

and decision-making are implicated in this process, relatively straightforward policy tools, such as stronger public pension schemes, may help to reduce inequality and poverty among the elderly. This is especially relevant given the dramatic shift away from defined benefit retirement plans towards options that give individuals greater financial autonomy (Poterba and Wise, 1998).

To measure ability endowments for human capital, we follow recent advances in behavioral genetics that have led to the identification of specific genetic markers that predict educational attainment. Results from the state of the art in this literature (Okbay et al., 2016) allow for the construction of a *polygenic score*—an index of these genetic markers—that robustly predicts education in the HRS (Papageorge and Thom, 2016). To explain the advantage of using this observed measure of genetic ability, it is useful to contrast our work with earlier studies linking ability endowments to economic behaviors that affect wealth. There are three approaches. One way is to use proxies for ability, such as cognitive test scores. However, since test scores are responsive to parental investments, their use may misattribute the importance of household resources to individual abilities. Another approach is to treat persistent factors as unobserved heterogeneity. While much of the work on persistent returns has done this, such an approach makes it difficult to understand the mechanisms that drive persistence.

More closely related to what we do, results from twins studies have shown that genetic factors play a non-trivial role in explaining saving behavior and portfolio decisions (Cronqvist and Siegel, 2014, 2015; Cesarini et al., 2010).[2] However, it is generally difficult to learn about how specific genetic factors operate using twins data. For one, twins studies do not identify which particular markers matter for wealth, nor do they allow one to measure these genetic endowments at the individual-level. This means that analysis of their function is limited to variance decompositions related to mechanisms that are measured in existing twins data sets. Moreover, while testing hypotheses about specific mechanisms is conceptually possible using information on twins, in practice it requires large amounts of data to permit stratification by each potential mediating factor.[3] Given that they are observable, the development of polygenic scores helps to overcome these difficulties.

We present three main sets of findings. Our first results establish a robust relationship

[2]For example, using the Swedish Twin Registry, Cesarini et al. (2010) demonstrate that about 25% of individual variation in portfolio risk is attributable to genetic variation while Cronqvist and Siegel (2015) show that 35% of variation in propensity to save has a genetic basis. It is worth mentioning, however, that these estimates may be biased upward if identical twins face more similar family environments than do non-identical twins (Fagereng et al., 2015).

[3]Variance decomposition exercises such as twins studies treat genes as unobserved factors. Learning about interactions between observed and unobserved factors is generally difficult and rests on modeling assumptions.

between the polygenic score and wealth. A one-standard-deviation increase in the score is associated with a 24 percent increase in household wealth (approximately $117,000 in levels). Education and household income account for a little more than half of this association. The relationship is only slightly moderated when we control for financial bequests, which could capture how individuals with higher polygenic scores have parents who are more highly educated and therefore wealthier. Despite these controls, the polygenic score continues to exhibit a strong and economically large relationship with household wealth. This leads us to explore how these genetic factors relate to investment decisions. A classic example of such a decision is stock market participation (Vissing-Jørgensen, 2002). We demonstrate that the polygenic score predicts stock market participation even after we have controlled for education, wealth, and wages, and that stock ownership plays an important role in mediating the relationship between genetic endowments and wealth. Measures of risk aversion available in the HRS do not appear to explain this relationship.[4] Because stock market participation affects returns to wealth, these results link the genetic endowments captured by the polygenic score to the persistent returns described in Benhabib (et al 2011).

Having shown evidence that at least part of the genetic gradient in wealth can be explained by financial decisions, our second set of results exploits rich data on subjective macroeconomic expectations to investigate one possible underlying mechanism. We show that the polygenic score is associated with expectations that are objectively more correct (Dominitz and Manski, 2007).[5] Furthermore, a lower score is associated with beliefs about the economy that are heaped on probabilities of 0% or 100% (a phenomenon we refer to as "extreme beliefs"), which suggests difficulty with probabilistic thinking. Importantly, these extreme beliefs do not appear to solely reflect respondent confusion or measurement error; households that report extreme beliefs make choices that are consistent with their beliefs. For example, households that report a 100% probability of a stock market increase over the following year are 20 percentage points more likely to invest in stocks than those that report a 0% probability.[6] We do not claim that this is the sole mechanism linking genetic ability,

[4]Previous research suggests that risk aversion has a genetic basis (Cesarini et al., 2009). Unfortunately, a measure of risk aversion is available only for a subsample of individuals in the HRS. Controlling for this measure does not substantially attenuate the genetic gradient, but we may be underpowered to detect meaningful differences.

[5]Hurd (2009) provides a review of subjective probabilities reported in household surveys such as the HRS. A number of researchers have used the HRS to study cognition, probabilistic thinking and investment decisions (Lillard and Willis, 2001; Kézdi and Willis, 2009, 2003). Another set of related studies focuses on cognitive decline and retirement decisions (Rohwedder and Willis, 2010; Kézdi and Willis, 2013; Delavande et al., 2006; Delavande, Rohwedder, and Willis, 2008).

[6]Lillard and Willis (2001) also recognize that focal point beliefs are predictive of investment behavior. This highlights an important distinction: individuals who report extreme beliefs do not necessarily make irrational choices at random, but rather make potentially optimal choices based on incorrect beliefs.

decision-making, and wealth. However, the results on beliefs offer a clear channel driving some individuals to make consistently better financial decisions and to access higher returns to wealth.

Our findings suggest that the persistent heterogeneity described in Benhabib, Bisin, and Zhu (2011) may at least partially operate through genetic endowments linked to human capital accumulation, financial decision-making, and the formation of expectations. Our analysis both complements and contrasts with the findings of Cronqvist and Siegel (2015), who use twins data to study a genetic basis for *savings* behavior. They conclude that genetic factors related to savings and wealth may operate through time preference and self control because of genetic correlations between savings, smoking, and obesity. Though our results confirm a genetic basis for heterogeneity in wealth, we appear to identify distinct mechanisms. To highlight this point, we consider additional polygenic scores developed to predict BMI and smoking (cigarettes per day). We find that all three polygenic scores independently predict wealth. In particular, we demonstrate that the polygenic scores for BMI and smoking are negatively associated with wealth, confirming the results in Cronqvist and Siegel (2015).[7] Nevertheless, we find that the polygenic score for education continues to predict wealth even after we condition on these other polygenic scores. Moreover, we do not find any robust associations between the scores for BMI or smoking and expectations. Cronqvist and Siegel (2015) find no overall genetic correlation between *savings* and education. Taken together, these results suggest that the genetic endowments linked to education, which are the primary focus of our study, affect wealth through mechanisms that are not directly tied to savings behavior.

Our third set of findings considers policy implications. If genetic ability affects the quality of financial decision-making, policies that grant individuals greater discretion in how they invest and manage financial decisions could exacerbate ability-based inequality. Comparing wealth outcomes of individuals with and without defined benefit pensions offers some evidence on this. In particular, we show that the gene-wealth gradient is much stronger for people who do not have such pensions as a source of retirement income. Our interpretation is that lower genetic scores are more problematic when individuals are tasked with making their own saving and investment decisions. Such evidence highlights the possible costs of institutional changes that increase autonomy in financial decision-making, such as the shift away from defined benefit pensions towards defined contribution plans. Correspondingly, this suggests that some groups may benefit from institutions that make financial decisions on their behalf. Promoting better financial decisions represents an alternative to such pater-

[7] Indeed, we believe this cross-validation demonstrates that twin studies and studies using polygenic scores yield similar patterns.

nalism. However, this is only viable if the genetic gradients we describe are easily modified by household resources and other environmental factors. To explore this possibility, we assess whether individuals with low genetic scores who are born into more advantageous environments are less likely to report "extreme beliefs". Although measures of childhood socioeconomic status (SES) have been shown to moderate the relationship between these genetic endowments and college completion in the HRS (Papageorge and Thom, 2016), we do not find that SES plays such a moderating role for wealth or the accuracy of subjective expectations.

The remainder of this paper is organized as follows. Section 2 provides some background on the genetic index used in this paper. Section 3 introduces the data used in this project, paying particular attention to the construction of wealth measures. Section 4 studies the relationship between the polygenic score and wealth. Section 5 discusses the polygenic score, financial decision-making, beliefs about the economy and probabilistic thinking. Section 6 discusses some policy implications of our results. Section 7 concludes.

2 Genetic Data and GWAS

We measure labor market ability using a polygenic score that predicts educational attainment. Since economists are only just beginning to use genetic data, we first provide a brief introduction to the genetic data we use, as well as its advantages and shortcomings. While we only present some main ideas here, Appendix A offers much more detail, and a host of references for the interested reader.

The first point concerns problems linking genetic data to economic behavior. An individual's genome consists of 23 pairs of chromosomes, one from each parent. Each chromosome can be divided into sections that are functionally related, called *genes*. Each gene is comprised of millions of *nucleotide pairs* — these are the "rungs on the ladder" in illustrations of our DNA — and such pairs can take only one of two values.[8] Further, humans only differ from one another in a few million of these pairs (less than 1% of the total). The pairs in which humans may differ are called *single nucleotide polymorphisms*, or SNPs (pronounced "snips").

Once scientists could observe these SNPs (when the human genome was sequenced about 15 years ago) researchers began to link specific SNPs to physical characteristics (e.g., hair

[8] Each pair can be an adenine-thymine (AT) pair or a guanine-cytosine (GC) pair. While each rung will be one of these two molecules, the rungs might differ in terms of their relationship to the sides or rails of the ladder. That is, we might have an AT molecule or a TA molecule, where different ends of the rung are connected to different sides of the ladder.

color), but also to behavior (e.g., smoking). However, with millions of SNPs, it is not clear how to identify which SNPs are relevant for a particular trait. An initial remedy was to use so-called "candidate genes", identified by theories about biological processes likely to be important for the behavior or outcomes of interest. A multiple-hypothesis testing problem arose, however, given the sheer number of possible candidates, even within a particular gene. This was such a problem that many encouraging results turned out to be false positives (Hewitt, 2012; Benjamin et al., 2012).

These challenges led behavioral geneticists to a new approach, known as genome-wide association studies (or GWAS). In a GWAS, researchers embrace an atheoretical approach and test each SNP individually for a relationship with the outcome of interest. Stringent controls are applied to account for multiple hypothesis testing. Essentially, all SNPs are regressed one by one, along with a set of essential controls. The GWAS revolution has produced a number of robust, credible results, including the discovery of the most well known genetic variant associated with obesity and several markers associated with smoking (Bierut, 2010; Thorgeirsson et al., 2010). Once a GWAS generates a series of coefficients associated with individual markers, these coefficients can be used to construct genetic indices called *polygenic scores*. These scores are typically linear combinations of individual markers. The Appendix provides considerably more detail on how this is done.

Our measure of genetic ability comes directly from a series of landmark GWAS discoveries that have identified some of the first direct associations between specific SNPs and educational attainment (Rietveld et al., 2013; Okbay et al., 2016). After documenting the first GWAS for education (Rietveld et al., 2013), the Social Science and Genetics Consortium recently extended their analysis to perform an educational attainment GWAS with an unprecedented sample size of 293,723 (Okbay et al., 2016). Our genetic measure is the polygenic score developed in this follow-up study, which combines all genotyped SNPs. We refer to this measure as the *EA Score*, indicating that "educational attainment" is the trait of interest. In recent work, Papageorge and Thom (2016) show that the EA Score predicts labor market outcomes independently of education, including wages and retirement.

We conclude this section by briefly discussing three important caveats and points of clarification. First, it is important to note that the genetic variants used in the construction of this genetic score are not located on sex chromosomes. For this reason, the distribution of these variants should be identical across men and women. Second, we do not claim to estimate *causal effects* of particular genetic variants. Any gene-outcome association that we observe in general reflects a combination of a direct effect and an indirect effect operating through the environments that parents make for their children. Parents with advantageous genetic endowments (which they pass on to their children) are more likely to have the resources or capacity to create better environments. Even so, an individual's genetic make-up is not *changed* by human capital investments. In contrast, IQ and other cognitive test scores are subject to the critique that they reflect environmental factors, such as earlier human capital investments. For example, Bharadwaj, Løken, and Neilson (2013) find that variation in health care received by newborns has an impact on academic achievement years later.[9] Genetic indices are not subject to this critique because they are fixed at conception.

A final caveat to our use of genetic data is that it may misrepresent ability. By aggregating a number of genetic variants into a single score, we are implicitly assuming that these factors work together in determining a single scalar value. However, a growing literature suggests that ability is multi-dimensional. There may be distinct cognitive abilities (e.g., mathematical ability or facility with language), each possessing different associations with economic outcomes (Willis and Rosen, 1979; Heckman, 1995; Cawley et al., 1997). Ability may also encompass not just cognition, but non-cognitive factors as well (Heckman and Rubinstein, 2001).[10] On this point, we are tied to the state of the art in genetics. We do not yet have the tools to credibly determine whether the individual genetic markers that make up the score contribute to distinct abilities.

3 The HRS Sample, the Polygenic Score and Wealth

In this section, we introduce the data set we use to examine how genetic ability endowments relate to wealth. Section 3.1 provides details on how we construct our sample and provides basic summary statistics. Section 3.2 provides details on our construction of household wealth.

3.1 Sample Construction

The Health and Retirement Study (HRS) is a longitudinal panel study that follows Americans over age 50 and their spouses. Surveys began in 1992 and occur every two years. The HRS collected genetic samples from 15,680 individuals over the course of three waves (2006, 2008, 2010). Genetic data from the 2010 wave have not yet been publicly released, so our sample

[9]Even birth weight, another proxy of innate endowments that has been used in prior literature, is not immune to this critique as it reflects *in utero* investments, e.g., mothers' smoking behavior (Lien and Evans, 2005), exposure to pollutants (Currie, Neidell, and Schmieder, 2009) stress during pregnancy (Camacho, 2008; Currie and Rossin-Slater, 2013) or mothers' own health (Costa, 1998). See also Aizer and Currie (2014) for a recent discussion.

[10]On multidimensionality, Willis and Rosen (1979) emphasize manual skill, which they distinguish from academic skill.

only includes individuals genotyped in 2006 and 2008.[11] Individuals in the genotyped sample tend to be born in younger birth cohorts because survival until at least 2006 is required for inclusion. Moreover, women and individuals with more education were more likely to agree to the collection of genetic data.

Our main analysis sample includes all genetically European financial respondents born before 1965 with non-missing genetic, education, and household wealth data.[12] We restrict the sample to European-Americans because the polygenic score we use was discovered in a sample consisting solely of genetic Europeans. We further restrict our sample to include only retired households in years 1996, 1998, and 2002-2010.[13] This restriction is aims to balance concerns about measurement error in wealth with concerns about selection biases that arise if we drop too many observations from our analysis. Further details on wealth data, including measurement problems, are found in the following section.[14]

The resulting analytic sample includes 4,433 financial respondents, with responses supplied for an average of seven waves.[15] Table 1 provides basic descriptives on demographic and educational variables. The mean level of educational attainment is about 13 years, with 21% of the sample failing to graduate from high school or obtain a GED and about 22% of the sample earning at least a four year college degree. Roughly 42% of the sample is male. For financial respondents in our analytic sample, Figure 1 plots the sample (kernel-smoothed) density of the EA Score variable, the genetic index score we use for our analysis. Values of the score have been demeaned and re-scaled to measure standard deviations relative to the mean. Figure 1 suggests that the distribution of the EA Score is approximately normal.

3.2 Household Wealth

The HRS contains rich and detailed information on household wealth. Unfortunately, data related to household retirement wealth and stock market participation pose various challenges. Values of defined contribution plans from previous jobs are not asked in every wave;

[11] Release of genetic data from the 2010 wave is imminent. Thereafter, the genetic score variable must be re-computed for the larger sample. At that point, it should be straightforward to extend our econometric analysis to include the larger sample.

[12] As part of the genetic data release, the HRS also released a file flagging 8,652 individuals as being of European descent based on their genes.

[13] By retired, we mean no member of the household is currently employed.

[14] In further robustness checks available from the authors we demonstrate that our results are robust to (and in many cases stronger) using alternative data samples. Therefore, we are confident that our main results relating ability to wealth are not driven by our choice of subsample.

[15] In subsequent analyses, sample sizes fluctuate depending on "missingness" of data. In particular, we obtain larger sample sizes for variables for which we have repeated observations not only from the financial respondent, but also from other members of the household. Moreover, when data are collected for small numbers of individuals (e.g., special modules), we include as many observations as possible.

stock allocations in defined contribution plans are only asked in certain waves, and only for plans associated with the current employer; expected defined benefit pension income is also asked only of plans at the current employer. In some cases, such issues may be relatively unimportant. However, because this paper studies heterogeneity in wealth for elderly households, having a complete picture of the households' retirement assets is of fundamental importance. While some data issues have no hope of being resolved, our sample comprises households for whom wealth data is most likely to be both accurate and comprehensive.

Our measure of *total wealth* in 2010 dollars is designed to encompass all components of household wealth. Our data includes the present value of all pension, annuity, and social security income, which come from the RAND HRS income files, as well as the net value of housing (including primary and secondary residences as well as investment property), the net value of private businesses, all financial assets including cash, checking accounts, savings accounts, CDs, stocks and stock mutual funds, bonds and bond mutual funds, trusts, and other financial assets, less the net value of non-housing debt, each of which we derive from the RAND HRS wealth files.[16] Further, we include the account value of all defined-contribution retirement plans. We exclude from our wealth measure values of transportation and insurance.[17]

We note that our measure of wealth includes both marketable securities, such as stocks which can be easily sold at publicly available prices, and non-marketable assets such as social security income. Our measure of wealth is therefore intended to capture the degree of financial security of the household, rather than the market value of household assets. In results available from the authors, we show that main results hold if we limit attention to household *financial wealth* (total wealth excluding retirement income and housing), which can be interpreted as the market value of households' salable financial assets. Further details on wealth data, reasons for possible mis-measurement and possible alternative subsamples are found in Appendix B.

Figure 2 shows the unconditional distribution of wealth for observations in our analytic sample. Notice that the distribution is right-skewed, which is consistent with a relatively small number of individuals who report very high levels of wealth. Figure 3 shows that the distribution of log wealth is somewhat more normally distributed.[18]

Table 2 shows the 10th, 25th, 50th 75th and 90th percentiles of our various wealth mea-

[16] We follow Yogo (2016) and assume a 1.5% guaranteed rate of return, discounted by the probability of death in each year conditional on age, cohort and gender as determined by the Social Security life tables.

[17] In principle, expected defined benefit income is knowable in all years between 1992 and 2012 for working respondents, but only for the current job.

[18] Recall from our discussion in Section 2 that the polygenic score does not reflect variants on sex chromosomes, so its distribution should be identical in men and women.

sures, as well as the average. This table contains data on all 15,517 person-year observations with non-missing wealth data. The average age is 73, ranging from 53-101.[19] The first row of Table 2 shows winzorised total wealth, which includes housing and pensions. The average for our sample is about $540,000. However, the median individual has total wealth of roughly $284,000, which is substantially lower. Again, this is due to high levels of wealth among individuals in the upper tail. For example, wealth at the 10th percentile is about $30,000; at the 90th percentile, wealth is a little over $1,230,000.

Rows two through four of Table 2 show wealth excluding the values of housing and pension income. A few interesting patterns emerge. First, housing makes up a larger portion of total wealth at the lower end of the distribution. For example, at the 10th percentile housing wealth is more than half of total wealth, whereas it accounts for only about one-fifth of total wealth at the 90th percentile. A similar result is found for pension wealth. In fact, for individuals at the 10th percentile, housing and pensions comprise the entirety of household wealth.

Table 3 shows the median, 75th and 90th percentiles of the individual components of total wealth. At the median, the table confirms that nearly all wealth is in the form of pensions and housing. The 75th percentile includes other sources of wealth, including IRA's, stock holdings, cash and CDs. At the 90th percentile, wealth is further diversified, including items such as secondary homes and real estate.

4 Genes, Wealth and Financial Decisions

In this section, we study the relationship between genetic labor market ability (as measured by the polygenic score), wealth and financial decisions. In Section 4.1 we establish that the relationship between the EA Score and household wealth is substantial. Controlling for factors such as education, bequests and household income reduces the size of the gradient by about half. In Section 4.2, we assess whether stock market participation can further explain the relationship between labor market genetic ability and wealth. Controlling for stock market participation accounts for an additional third of the size of the genetic gradient in wealth. In Section 4.3, we examine whether risk preferences explain the relationship between the EA Score and stock market participation and show that they do not. Finally, in Section 4.4, we demonstrate that our main results are robust when we incorporate

[19]This age range raises the possibility that our analysis not only captures the relationship between the EA Score and wealth accumulation, but also reflects the running down of assets as individuals age post-retirement. However, 75% of the sample is under aged 80. Basic associations between the EA Score and wealth remain the same if we restrict attention to this younger subsample.

spousal ability. Taken together, our results suggest a robust and economically meaningful relationship between household genetic ability and wealth. However, this link does not solely operate mechanically through higher earnings, but also appears to reflect how individuals make financial decisions. In Section 5 we examine preferences and expectations — individual characteristics that affect financial decision-making.

4.1 The Polygenic Score, Household Wealth and Earnings

Figure 4 presents the stylized fact that motivates this paper. The nonparametric, unconditional relationship between the polygenic score of the financial respondent and household wealth, shown by a Lowess plot, is both positive and economically substantial. In Table 4 we regress wealth on the score and various sets of control variables. Unless otherwise indicated, all specifications throughout the paper include the following basic controls: the first ten principle components of the genetic data, a full set of birth year dummies, a full set of age dummies, a full set of calendar year dummies, a male dummy, and interactions between the male and age dummies and the male and birth year dummies. Panel A presents results for the log of wealth, and Panel B presents results for wealth in levels. The results in Column (1) indicate that a one-standard-deviation increase in genetic ability is associated with 24% higher total wealth (or about $117,000 in levels).

Since the score measures genetic endowments that predict educational attainment, it is natural to examine how much of this gradient can be explained by education. Column (2) adds controls for the financial respondent's years of schooling and degree, which reduces the coefficient by more than half; observed educational investments unsurprisingly play a large role in mediating this relationship. In Column (3), adding controls for parental education barely reduces the gradient, as might be expected given the strong inter-generational persistence in education. Regardless of education controls, the coefficients in the third column show that a strong and significant association remains. After controlling for own and parental education, a one-standard-deviation increase in the EA Score is associated with a 10% increase in wealth (or about $54,000).

We next explore mechanisms that might explain the strong positive association between the polygenic score and wealth, even after controlling for education. Since the score is directly related to human capital accumulation, perhaps the most obvious channel is performance in the labor market. Indeed, Papageorge and Thom (2016) demonstrate that the score is associated with higher wages after controlling for education and family background. Column (1) of Table 5 presents the association between the score and log wealth, now restricting to the sample with non-missing data on income and the other mechanism variables considered here (coeff. = 0.12).[20] In Column (2) of Table 5, we additionally control for the average log household income (averaged over years with non-missing household labor income). Though household income predicts higher household wealth as expected, the coefficient on the EA Score remains nearly the same. This is not terribly surprising, since Papageorge and Thom (2016) find a modest average relationship between the score and personal income, with substantial gradients only appearing for the college educated evident among more recent birth cohorts.[21]

Since individuals receive their genes from their parents, there will necessarily exist a high degree of correlation between their genetic score and the scores of their parents. High ability individuals with more education will tend to have more successful, high ability parents. This naturally suggests bequests and inheritance as a possible mechanism linking the polygenic score to household wealth. In Column (3) of Table 5, we add two separate controls for inheritances. First, we add the log of cumulative value of inheritances received in the current year (plus one). Second, we add a binary indicator for whether the individual ever receives inheritances in the HRS data.[22] Adding these inheritance variables reduces coefficient on the EA Score only modestly from 0.121 to 0.106.

Besides influencing the earnings of an individual (and their parents), genetic ability could affect household wealth by altering the decisions that households make with their saved earnings and other financial resources. Entrepreneurship or business ownership represents one investment choice that could drive systematic differences in wealth (Quadrini, 2000). In Column (4) of Table 5, we add an additional control for whether any member of the household has ever been observed owning a business in the HRS. This is the case for approximately 35.5% of individuals in the sample. This measure of business ownership is associated with substantially higher wealth (approx. 30 percent), but its inclusion does little to moderate the coefficient on the EA Score, which declines from 0.106 to 0.097.

Together, our main findings from this section indicate that about half of the genetic gradient in wealth can be explained by education and earnings. In other words, the relationship between labor market ability and wealth is to some degree a mechanical consequence of labor market outcomes: people with higher earnings and access to bequests accumulate

[20]Note that this coefficient is slightly higher than the coefficient delivered by the specification in Column (3) of Table 4, which used exactly the same control set. This is due to the restriction in the sample to those with non-missing values of the income, inheritance, and stock market participation variables examined in this section.

[21]After conditioning on our standard control set (excluding own and parents' education), the incremental R^2 of the EA Score is 2.2% Once we control for own and parents' education, it falls to 0.4% For comparison, the EA Score predicts 6.6% of the variation in education outcomes once we condition on our standard control set excluding parents' education.

[22]43.5% of individuals in the sample report an inheritance.

larger amounts of wealth. Nevertheless, much remains to be explained. We now ask if we can further explain the genetic gradient in wealth by considering financial decisions — in particular, stock market participation.

4.2 Examining the Role of Stock Market Participation

One of the most prominent financial decisions facing a household is how to allocate their wealth across different asset classes. Indeed, an enormous literature explores the portfolio choice problem and stock market participation in particular (Van Rooij, Lusardi, and Alessie, 2011). Returning to Table 5, in Column (5) we add an indicator variable for any stock market participation as a regressor in our specification. Accounting for stock market participation reduces the coefficient on the EA Score by more than one-third, from 0.099 to 0.062. This suggests that portfolio choice decisions, and stock market participation in particular, may represent a critical channel linking genetic ability to household wealth.

To learn more about the relationship between genetic ability and stock market participation, Table 6 estimates specifications in which the dependent variable is a dummy for whether or not the financial respondent's household owns any stocks. Column (1) indicates that a one-standard-deviation increase in the EA Score is associated with a 3.7 percentage point increase in the probability of stock market participation. Two explanations could rationalize this pattern. First, this could simply reflect the fact that individuals with higher polygenic scores are wealthier, and wealthier people tend to invest more of their wealth in risky assets such as stocks. Alternately, this genetic gradient could reflect differences in the portfolio choices that people make for a given level of wealth. In Column (2) 6, we add the lagged (last wave) values of the log of household wealth as a control variable. This is indeed strongly associated with stock market participation, and its inclusion reduces the coefficient on the EA Score to 0.028, cutting the association by about one fourth. When the average log household income is included in Column (3), the coefficient remains the same. Our conclusion is that the genetic score appears to be associated with stock market participation, even controlling for current wealth and the average of past wages.

4.3 Risk Preferences

Given the importance of stock market participation in accounting for the genetic gradient, a natural hypothesis is that the endowments captured by the EA Score may operate through risk preferences. To examine this mechanism, we use survey items in the HRS that pose hypothetical questions about a choice between guaranteed total family income or a gamble that might result in a permanent increase or decrease in total family income. Specifically, respondents are asked to choose between two jobs: "The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by X." The series replaces X with a set of possible income losses: "10 percent, twenty percent, a third, half, seventy-five percent." We create a dummy variable for risk aversion which takes a value of one if an individual always responds with a preference for the job that guarantees current income over the job that might either double income or result in a 10 percent loss. This response indicates the highest degree of risk aversion permitted with this set of questions, and approximately 32 percent of financial respondents in our basic sample always respond this way.[23]

We explore the relationship between risk aversion, the EA Score and wealth in Table 7. In Column (1), the dependent variable is our binary indicator for risk aversion. We find a weak negative association between the EA Score and risk aversion — a one-standard-deviation increase in the score is associated with a reduction in the probability of a risk averse response by 1.8 percentage points. However, this association is only marginally significant ($p < 0.10$). In Columns (2)-(3), we revisit our basic specification with the log of total wealth as the dependent variable. Since we only observe the risk aversion measure for a subset of our baseline sample, we first re-estimate our basic specification (Column (3) from Table 4) on the sample with non-missing risk aversion measures. We find a coefficient on the EA Score of 0.124. In Column (3), we add our risk aversion dummy. Our measure of risk aversion is informative; risk averse individuals are estimated to have approximately 17 percent less wealth, and this association is highly significant. However, including this measure of risk aversion only slightly reduces the coefficient on the EA Score, from 0.124 to 0.121. This suggests that risk preferences, at least as captured by this measure, do not play a major role in explaining the genetic gradient in wealth.

Columns (4)-(5) of Table 7 consider stock ownership. As with total wealth, we first re-estimate our basic specification using the risk aversion sample. In Column (4) we find that a one-standard-deviation increase in the EA score is associated with a 3.5 percentage point reduction in the likelihood of stock ownership. When we add our risk aversion measure in Column (5), this association falls only slightly, to 3.4 percentage points. However, our risk aversion measure is indeed strongly associated with stock ownership; the probability of stock ownership is 7.3 percentage points less likely among risk averse individuals, and this association is highly significant. Taken together, the results in Table 7 indicate a weak

[23]One could imagine creating several measures of risk aversion based on this series of questions. For example, one could code individuals as being risk averse based on a different threshold (e.g. taking the guaranteed salary when compared to gamble with a possible loss of one third or more). In results available from the authors, we show that findings in Table 7 are robust to alternate cutoffs.

negative association between the EA Score and risk aversion that explains, at best, only a tiny portion of the relationship between the EA Score and wealth.

4.4 Non-Financial Respondent Ability and Household Wealth

Before turning to a more detailed discussion of the EA Score and financial decision-making, we briefly assess the robustness of our results if we incorporate the ability of other members of the household. Our analysis until now has only used the ability of the "financial respondent" (FR). Doing so ignores the possibility that household wealth may also be a function of the ability of the financial respondent's spouse, deemed the "non-financial respondent" (NFR). We demonstrate that NRF ability is predictive of wealth even after controlling for FR ability. However, we also demonstrate that our results are qualitatively similar if we instead consider maximum household ability.

In Panel A of Table 8, we show that the scores of both the NFR and FR independently predict wealth if we include both in a regression with log wealth as the outcome variable. In Column (1), we restrict the sample to the set of households where both FR and NFR EA Scores are available and we regress log wealth (plus one) on the EA Score of the FR along with the standard set of controls and obtain a coefficient of 0.097. In Column (2), if we include EA Score of the NFR, we find that the coefficient on the FR EA Score falls modestly to 0.092 and that the coefficient on NFR EA Score is 0.078. In Columns (3) and (4), we repeat the exercise with stock market participation as the outcome variable. If we do not include NFR ability, the coefficient on FR ability is 0.022. This falls to 0.019 once we control for NFR EA Score. Surprisingly, the coefficient on NFR labor market ability is larger, estimated at 0.034.

Together, the results in Panel A of Table 8 provide support for the idea that basic patterns we have shown until now remain intact if we include NFR EA Score as an additional variable. However, these results also seem to suggest that the FR EA Score may not sufficiently capture the ability endowments that are relevant for household wealth outcomes. That is, there may be alternative ways to incorporate the labor market ability of both spouses. This naturally raises the question of how the ability endowments of each spouse combine to produce joint household outcomes. The endowments of the FR and the NFR might substitute for one another if an individual's high ability can compensate for the low ability of their spouse. Alternately, spousal abilities might be complementary if high and low ability spouses have to reach compromise positions on financial decisions.

We defer a full analysis of this topic for future work. However, Panel B of Table 8 provides some suggestive evidence that a high ability spouse can compensate for the low ability of a financial respondent. Here we consider as a regressor not the FR EA Score, but the maximum of the FR and NFR scores. In Column (1), we replicate our baseline wealth regression, restricting to households with two genotyped spouses (EA Score coeff = 0.097). In Column (2) we instead measure household ability with the maximum score variable, and the estimated genetic gradient increases by over one third (coeff = 0.147). Similar results hold in Columns (3)-(4), where we repeat this exercise for stock market participation.

5 Genes, Financial Literacy and Expectations

The preceding results suggest that i) there is a substantial relationship between the EA Score and household wealth, ii) this is not entirely or even primarily related to higher income, and iii) this appears to be at least partially mediated by stock market participation, which is a financial choice variable. This leads us to investigate whether individuals with different levels of the polygenic score differ in terms of how they think about financial decisions. In Section 5.1, we document a relationship between genetic ability and financial literacy. In Section 5.2, we examine genetic ability and subjective beliefs about macroeconomic outcomes. Individuals with higher values of the EA Score report subjective expectations that are closer to objective probabilities, and are less likely to report "extreme" beliefs that take some outcomes as certainties. In Section 5.3, we show that these self-reported beliefs, though not incentivized, are indeed linked to wealth and financial outcomes. For example, individuals with lower genetic scores are more likely to predict excessively high probabilities of low stock market returns and are subsequently more likely to be observed avoiding the stock market. Finally, in Section 5.4 we explore whether the EA Score studied here is related to BMI or smoking, which earlier research has linked to savings (Cronqvist and Siegel, 2015).

5.1 The Polygenic Score and Financial Literacy

An obvious theoretical starting point is financial literacy (Delavande, Rohwedder, and Willis, 2008). Do individuals with higher EA Scores exhibit greater sophistication in their understanding of financial choices? Fortunately, the HRS data contain a number of questions that directly assess an individual's financial literacy. Unfortunately, these are asked in a small module in the 2010 wave, which leaves us with less than 700 genotyped respondents for these questions. The 2010 module asks three basic financial literacy questions:

- **Compounding Interest:** "First, suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow – more than \$102, exactly \$102, or less than \$102?"

- **Real Interest Rate:** "Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than today, exactly the same as today, or less than today with the money in this account?"

- **Diversify Stocks:** "Do you think that the following statement is true or false: buying a single company stock usually provides a safer return than a stock mutual fund?"

Columns (1)-(3) of Table 9 present linear probability models explaining whether respondents correctly answered these questions as function of the EA Score and our standard set of controls.[24] The score is positively related to correctly answering the Real Interest Rate and Diversification questions, but only the coefficient for the Real Interest Rate questions is statistically significant (p-value < 0.05). In Column (4), the dependent variable is an indicator for whether the individual correctly answered all three questions. We find that a one-standard-deviation increase in the EA Score is associated with a 4 percentage point increase in the probability of correctly answering all questions (p-value <0.1). The financial literacy module also asks a separate question on whether creditors or debtors would be helped by inflation. This question is not asked to the individuals who answered the three questions listed above. Column (5) present estimates related to a correct answer on this question. The genetic association for this question is statistically significant and economically substantial: a one-standard-deviation increase in the score is associated with an 8 percentage point increase in the probability of a correct answer (p-value < 0.01).

Taken together, the results in Table 9 provide some evidence that individuals with higher values of the genetic score tend to be more financially literate. However, we again reiterate that these results should be interpreted cautiously since the financial literacy questions are available only for a small subset of the individuals in our main analysis sample. Thus, we may be underpowered to detect true effects with such relatively small samples.

5.2 The Polygenic Score and Financial Expectations

An important element of financial decision-making is an assessment of the risks and uncertainties associated with the macroeconomy and the payoffs to different possible financial choices. Despite the typical assumption of rational expectations, it has long been recognized that individuals may have trouble forming accurate beliefs about probabilistic outcomes

[24]Those responding that they "Don't Know" were coded as not responding correctly.

(Savage, 1954; Kahneman and Tversky, 1972). Recent literature examines the role of subjective expectations in economic decisions such as human capital investments (Wiswall and Zafar, 2015) and stock market participation (Arrondel, Calvo Pardo, and Tas, 2014). Related, a number of papers have used HRS data to study the relationship between beliefs and investment behavior (Hudomiet, Kézdi, and Willis, 2011).

Here we investigate whether the EA Score is associated with differences in the beliefs and expectations about objective macro events that are relevant for financial choices. The HRS data are uniquely well-suited for this analysis, since most respondents are repeatedly asked to provide subjective probabilities on a range of events. Individuals are asked to provide a probability on a scale of 0 to 100, for the following three events:

- **Stock Market Goes Up**: "By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Dow Jones Industrial Average will be worth more than they are today?"
- **Economic Depression**: "What do you think are the chances that the U.S.economy will experience a major depression sometime during the next 10 years or so?"
- **Double Digit Inflation**: "And how about the chances that the U.S. economy will experience double- digit inflation sometime during the next 10 years or so?"

The panels of Figure 5 present histograms of the responses, pooling all person-year observations across waves. Across all three questions, we see evidence of heaping at focal probabilities. Specifically, there are pronounced spikes at answers of 0, 50, and 100. It is essential to note that respondents are given specific instructions about the meaning of a response of 0 or 100. That is, they are told to supply these answers if they believe that there is "absolutely no chance" that the event will happen, or if it is "absolutely certain" to happen. All three of these macroeconomic events can be associated with objective probabilities that should be common knowledge in an economy with fully informed agents. Indeed, Hudomiet, Kézdi, and Willis (2011) also discuss such "focal point" beliefs, compare them to objective probabilities and recognize that they can drive behavior.

Our objective benchmark probability for the stock market going up in a single year is 71 percent, which is the average number of years of positive returns for the S&P 500 over the period 1992-2015. There is no common definition of an economic depression, but clearly this refers to an unusually severe period of economic contraction. We use data from the Federal Reserve Bank of Saint Louis on annual real GDP growth over the period 1948-2016, and define an unusually severe contraction as a year with growth less than or equal to

-0.73 percent, which is the 25th percentile of the distribution of growth rates for negative-growth years. Based on this metric, the unconditional probability of a severe contraction is 4.4 percent per year, which implies a 36 percent probability for such an event over a 10 year period. Finally, we note that the Bureau of Labor Statistics reports two years with double digit inflation (1980, 1981) over the period 1958-2015. This implies an approximate probability of 3.4 percent for double digit inflation in any year, or about a 29 percent chance for double digit inflation over a 10 year period.

Table 10 provides estimates of the association between the EA Score and individual beliefs about the probabilities of these macroeconomic events. Each panel presents results on a different expectation (Panel A: probability the stock market goes up, Panel B: probability of a depression, Panel C: probability of double digit inflation). Our first measure related to these expectations variables is the absolute value of the deviation between the respondent's probability and the objective probability. We regress this deviation on our standard controls and the EA Score in Column (1). For all three events, higher values of the polygenic score are associated with a statistically significant reduction in the deviation between the respondent's subjective probability and the objective probability. For example, in Column (1) of Panel A, the coefficient estimate of -0.453 suggests that a one-standard-deviation increase in the score is associated with a reduction in the deviation of about one half of a percentage point.

Columns (2)-(4) of Table examine binary outcomes indicating whether respondents answered with specific focal probabilities (0, 50, and 100, respectively). For all three events, we observe the same pattern of association: The EA Score is negatively associated with providing a subjective probability indicating complete certainty (0 or 100), and is largely uncorrelated with providing a focal probability of 50 percent. It should be noted that the magnitudes of these associations are substantial. For example, Column (2) of Panel B suggests that a one-standard-deviation increase in the EA Score is associated with a 0.5 percentage point reduction in the probability of believing there is 0 chance that the economy will suffer a major depression in the next 10 years. For the sake of comparison, 6.7 percent of individuals responded with a 0 belief for this item.

5.3 Linking Financial Beliefs to Wealth and Financial Decisions

The results of the preceding section suggest that individuals with lower genetic scores appear to report beliefs that are at odds with objective probabilities and, moreover, tend to heap on "focal" beliefs. One possibility is that these reported beliefs are not related to individual behavior in a meaningful way. For example, it may be that individuals with lower polygenic scores simply report numbers that are not reflective of their beliefs. This would suggest that their answers are more prone to measurement error, but do not necessarily imply greater difficulties with financial decisions.

In Table 11, we investigate how beliefs relate to behavior and wealth. In Column (1), we show that individuals who ever predict a 0 belief that the stock market will go up are significantly less wealthy by about 22 percentage points. Similar magnitudes are evident for individuals who report that they are certain there will be a recession or double-digit inflation. Interestingly, individuals who report they are certain that the stock market will go up report gains to their wealth of about 26%. In Column (2), we add the deviations discussed in the previous section and find that it is not only focal beliefs, but also deviations from the objective probabilities that matter, at least for the stock market and depression events. To investigate these patterns a bit further, Columns (3) and (4) relate beliefs to whether or not the individual owns stocks. Not surprisingly, individuals who are pessimistic are less likely to participate in the stock market, while individuals who are certain that the market will go up are more likely to participate in the stock market. This table suggests that expectations about the economy (including focal beliefs and beliefs that deviate from objective probabilities) predict individual financial decision-making (Hudomiet, Kézdi, and Willis, 2011). Notice, in particular, that individuals who are certain that the stock market will go up are substantially more likely to hold stocks (by about 10 percentage points). This is consistent with a world in which beliefs matter because they directly inform decisions.

5.4 Alternate Mechanisms

We have stressed expectations and information processing as two likely mechanisms linking genetic endowments for education with wealth accumulation. However, it is plausible that the EA Score also measures genetic factors that operate through other channels. In particular, Cronqvist and Siegel (2015) argue that the genetic endowments that drive savings and wealth may work through time preference and self-control. Using twin study methods, they demonstrate a genetic correlation between savings, obesity, and smoking. Given the role of time preference and self-control in governing health behaviors such as food intake or cigarette consumption, Cronqvist and Siegel (2015) suggest that these mechanisms are likely to also play an important role in the genetic basis for wealth. This raises the possibility that the EA Score also captures genes related to self-control and time preference, and that these traits provide a common genetic basis for education, smoking, BMI, and wealth. Alternately, the genes summarized in the EA Score may capture additional, distinct behavioral channels.

To assess these possibilities, we next turn our attention to additional polygenic scores that have been developed for BMI and smoking. These scores are constructed using the same methods used to develop the EA Score from Okbay et al. (2016). For obesity we use a polygenic score developed to predict BMI based on the estimates of Locke et al. (2015). For smoking, we use a score developed to predict the number of cigarettes smoked per day (CPD) at peak consumption, based on the estimates of Thorgeirsson et al. (2010). All scores have been standardized to have a mean of zero and a standard deviation of 1.

To begin, Panel A of Table 12 presents the simple correlations between the polygenic scores for education, BMI, and cigarettes per day. We find negative correlations between the EA Score and the other two polygenic scores. However, these correlations are relatively modest (-0.18 for BMI and -0.09 for cigarettes). We next explore the associations between the additional polygenic scores and wealth. In Panel B we add the BMI and smoking scores to some of our basic regressions for wealth and expectations accuracy. Column (1) adds these scores to our baseline specification explaining the log of total wealth (Column 3 of Table 4). The polygenic scores for BMI and CPD are both negatively correlated with wealth, consistent with the results from Cronqvist and Siegel (2015). A one-standard-deviation increase in the score for BMI is associated with approximately 6 percent lower wealth. This association is 5.3 percent for the CPD score. We believe that these results provide an important cross-validation of results in (Cronqvist and Siegel, 2015), which were generated using twin studies.

Next, we ask whether the BMI and CPD scores explain the relationship between the EA Score and wealth. Adding the BMI and CPD scores does moderates the relationship between the EA Score and wealth, but not substantially. The coefficient on the EA Score is still large and statistically significant (0.086 compared to 0.099 for the baseline). It is noteworthy that all three scores independently predict wealth. This could arise either because all three scores measure the same latent genetic endowments with error, or because they indeed reflect distinct genetic factors. To explore this further, we examine whether these other polygenic scores also predict extreme expectations. Columns (2)-(5) in Table 12 revisit our specifications on deviations from objective beliefs about the stock market, the risk of a depression, and the probability of double digit inflation (Column 1 from Table 10). While we continue to find a negative association between the EA Score and deviations from objective beliefs, we find no statistically significant relationships between the BMI and CPD scores and these deviations.

Taken together, the results from Table 12 suggest a nuanced story. Genes related to education, smoking, and obesity all appear to influence wealth. However, while the genetic endowments tied to education are associated with expectations and information processing, the endowments related to BMI and smoking seem to operate through distinct mechanisms.[25]

[25] In separate analyses available upon request, we replicate every regression in Table 10, adding in all three

The results presented here highlight the usefulness of polygenic scores and molecular genetic data for the study of population heterogeneity. The direct observation of polygenic scores allows one to extend genetic analysis to data sets like the HRS, which do not contain twins data, but do contain genetic data and rich information on variables such as beliefs that may not be accessible in twins studies. The results presented here have several implications for our understanding of heterogeneity in the wealth accumulation process. First, while twins studies have established the importance of genes for wealth, there appear to be multiple distinct genetic mechanisms. One source of genetic heterogeneity may work through time preference and self control, affecting savings along with health behaviors. We also find evidence for a separate source of heterogeneity — linked to human capital accumulation — which may work through information processing and expectations formation.[26] Because time preference and other savings-based mechanisms cannot explain persistent differences in *returns* to wealth, evidence for a genetic mechanism related to information and decision-making offers an important biological micro-foundation for the kind of heterogeneity described by Benhabib, Bisin, and Zhu (2011). Moreover, because this mechanism affects decision-making and not preferences, this raises the possibility that policies targeting information or assistance in financial choices could impact genetic inequality.

6 Evidence on Policy Implications

In this section, we turn to some policy implications of our results. The results of the preceding sections suggest that individuals with different levels of labor market ability differ in terms of financial literacy and the accuracy of their beliefs. Since these are important inputs for effective financial decision-making, individuals with lower levels of labor market ability might benefit if at least some of their financial decisions are managed by a third party. Defined benefit pensions (employer based pensions) offer one arrangement in which savings

polygenic scores. We continue to find statistically significant, negative associations between the EA Score and the likelihood that an individual provides an extreme probability in either direction for all three expectation outcomes. For the BMI and CPD scores, we find only marginally significant ($p < 0.10$) associations for two out of twelve associations, but otherwise find no statistically significant relationships between the scores and extreme beliefs. We do, however find that the BMI and CPD scores appear to be negatively associated with the probability of providing a subjective probability of exactly 0.50 for the stock market question (significant at the 0.05 and 0.10 levels, respectively). However, we note that we do not find the same pattern for the depression or inflation questions. We never find a statistically significant relationship between the EA Score and providing a subjective belief of 0.50 for any question.

[26] The HRS does feature an off-wave questionnaire on consumption that could be used to estimate savings behavior. Unfortunately, the sample sizes for this questionnaire are very small, preventing us from conducting a well-powered analysis of savings behavior. In results available upon request, find no significant relationship between the EA Score and the log of household consumption, controlling for the log of last year's labor income and our standard controls.

are effectively managed by an individual's employer (or a third party), and an individual receives a guaranteed stream of income without having to make investment decisions over the life-cycle. Of course, pension participation is not necessarily randomly assigned. In Table 13, we regress an indicator for defined benefit pension holdings onto the EA Score. We show that after including our standard battery of controls, there is no economically meaningful or statistically significant relationship between the EA Score and pension holdings. In Column (2), we assess the association between pensions and wealth and show that holding a pension is associated with a 37 percent increase in wealth. In Column (3) of Table 13, we ask if receiving a pension moderates the relationship between genes and wealth. In particular, aside from our general set of controls, the EA Score and a dummy for holding a pension, we also include an interaction between the EA Score and the pension dummy. The coefficient on the interaction is negative and significant, suggesting that pensions moderate the relationship between genetic ability and wealth.

Defined benefit pensions may reduce the genetic gradient by limiting autonomy in financial decision-making and protecting individuals from mistakes or incorrect beliefs. To test this particular account, we estimate an additional specification in Column (4) of Table 13 that uses information on (maximum) deviations from objective beliefs about the stock market, the risk of a depression, and the probability of double digit inflation (Column 1 from Table 10). We include these deviations as covariates and also interact them with the indicator for receiving a defined benefit pension. We continue to find significant negative relationships between each of these deviation variables and household wealth. However, we also find significant positive interactions between the defined benefit indicator and each of these deviations. This suggests that the consequences of inaccurate beliefs may be less pronounced for people faced with fewer investment decisions. When these interactions are included the interaction between the EA Score and the defined pension dummy shrinks in size (from -0.079 to -0.059) and becomes insignificant. While we cannot precisely estimate a difference in this coefficient across specifications, the results in Columns (3)-(4) are consistent with idea that the genetic endowments we study operate at least partly through expectations, and that disparities arising from these endowments can be mitigated in policy environments that demand fewer household decisions.

We also ask if high childhood SES mitigates the likelihood that individuals report extreme beliefs. Some individuals with disadvantageous ability endowments, but who are born into high-SES environments, may have access to resources and investments that would improve their ability to form expectations or to process information. We consider four retrospective childhood SES measures: whether or not the individual grew up in poverty; average income of father's occupation; whether the family ever moved or asked for help due to financial difficulties and whether the individual's father was ever unemployed for long periods of time.[27] In Table 14, we relate extreme beliefs to "high SES" for each measure. In Panel A, we consider the belief that the stock market will go up with 100% probability. High SES (not growing up in poverty) is associated with a lower probability of reporting an extreme belief. If the family moved or asked for help due to financial troubles, the individual is more likely to report an extreme belief. We ask whether there is an interaction between SES and a higher genetic score and find little evidence of such an effect. In other words, there is little evidence that high SES attenuates the relationship between a low polygenic score and extreme beliefs. We find similar results for extreme beliefs except for beliefs about double digit inflation and the income measures of childhood SES.[28]

The results on education provide some evidence that education many not be a solution to the problem that some ability endowments are associated with difficulties in probabilistic thinking. Of course, much research would need to be done to assess whether this is truly the case. Still, our results provide additional support for the idea that public pension schemes could mitigate inequality due to poor financial decision-making. These types of schemes have come under fire precisely because of the sound argument that individuals, acting on their own, could make better portfolio decisions than a public pension scheme, for example. The opposing argument is that poverty among the elderly, beyond being undesirable on ethical grounds, also imposes costs on society. If so, then allowing people to make autonomous financial decisions that lead to poverty creates an externality and it may be welfare enhancing to reduce autonomy.

7 Conclusion

This paper shows that the same genetic endowments that predict educational attainment and earnings are also associated with higher wealth. This could arise purely from an association between ability and earnings, as high earnings will mechanically generate high wealth. We show that controlling for education and earnings does indeed attenuate the genetic gradient in wealth, but only accounts for roughly one half of the association. Stock market participation accounts for one third of the remaining gradient in wealth, which suggests financial decision-making as another mechanism linking labor market ability and wealth. We show that those with a higher polygenic score perform better on standard financial literacy questions, and

[27]These four childhood SES measures are discussed at length and used in conjunction with the polygenic score in Papageorge and Thom (2016).

[28]Results remain the same under various sets of controls, i.e., whether or not we control for parents' or own education.

are less likely to report extreme beliefs about the economy, including the likelihood of stock market appreciation or a severe recession.

We also show evidence that childhood SES does not appear to modify the relationship between the polygenic score and beliefs. This is troubling as it suggests that reallocating resources for educational investments is not an easy solution for difficulties in financial decision-making. On the other hand, we show that the genetic gradient in wealth is weaker among individuals who have less autonomy in their financial decisions due to participation in traditional pension plans, while participation in pensions is not itself predicted by the polygenic score. Our findings suggest that policies that reduce autonomy in financial decision-making, such as public pension schemes, might play an important role in reducing wealth inequality. This is particularly important given our findings that the same ability endowments that predict low earnings also predict disadvantageous financial decision-making, which could further exacerbate wealth inequality among the elderly.

References

Aizer, Anna and Janet Currie. 2014. "The Intergenerational Transmission of Inequality: Maternal Disadvantage and Health at Birth." *Science* 344 (6186):856–861.

Arrondel, Luc, Hector F Calvo Pardo, and Derya Tas. 2014. "Subjective Return Expectations, Information and Stock Market Participation: Evidence from France." Mimeo, University of Southampton.

Bach, Stefan, Andreas Thiemann, and Aline Zucco. 2015. "The Top Tail of the Wealth Distribution in Germany, France, Spain, and Greece." DIW Berlin Discussion Paper.

Benhabib, Jess and Alberto Bisin. 2016. "Skewed Wealth Distributions: Theory and Empirics." NBER Working Paper.

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2015. "Wealth Distribution and Social Mobility in the US: A Quantitative Approach." NBER Working Paper.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents." *Econometrica* 79 (1):123–157.

Benjamin, Daniel J, David Cesarini, Christopher F Chabris, Edward L Glaeser, David I Laibson, Vilmundur Guðnason, Tamara B Harris, Lenore J Launer, Shaun Purcell, Albert Vernon Smith et al. 2012. "The Promises and Pitfalls of Genoeconomics." *Annual Review of Economics* 4:627–662.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review* 103 (5):1862–1891.

Bierut, Laura Jean. 2010. "Convergence of Genetic Findings for Nicotine Dependence and Smoking Related Diseases with Chromosome 15q24-25." *Trends in Pharmacological Sciences* 31 (1):46–51.

Camacho, Adriana. 2008. "Stress and Birth Weight: Evidence from Terrorist Attacks." *American Economic Review* :511–515.

Cawley, John, Karen Conneely, James Heckman, and Edward Vytlacil. 1997. "Cognitive Ability, Wages, and Meritocracy." In *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*, edited by Bernie Devlin, StephenE. Fienberg, Daniel P. Resnick, and Kathryn Roeder. Springer New York, 179–192.

Cesarini, David, Christopher T Dawes, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. 2009. "Genetic Variation in Preferences for Giving and Risk Taking." *Quarterly Journal of Economics* 124 (2):809–842.

Cesarini, David, Magnus Johannesson, Paul Lichtenstein, Örjan Sandewall, and Björn Wallace. 2010. "Genetic Variation in Financial Decision-Making." *The Journal of Finance* 65 (5):1725–1754.

Costa, Dora L. 1998. "Unequal at Birth: A Long-Term Comparison of Income and Birth Weight." *The Journal of Economic History* 58:987–1009.

Cronqvist, Henrik and Stephan Siegel. 2014. "The Genetics of Investment Biases." *Journal of Financial Economics* 113:215–234.

———. 2015. "The Origins of Savings Behavior." *Journal of Political Economy* 123 (1):123–169.

Currie, Janet, Matthew Neidell, and Johannes F Schmieder. 2009. "Air Pollution and Infant Health: Lessons from New Jersey." *Journal of Health Economics* 28 (3):688–703.

Currie, Janet and Maya Rossin-Slater. 2013. "Weathering the Storm: Hurricanes and Birth Outcomes." *Journal of Health Economics* 32 (3):487–503.

Delavande, Adeline, Michael Perry, Robert Willis et al. 2006. "Probabilistic Thinking and Early Social Security Claiming." Michigan Retirement Research Center Working Paper.

Delavande, Adeline, Susann Rohwedder, and Robert J Willis. 2008. "Preparation for Retirement, Financial Literacy and Cognitive Resources." Michigan Retirement Research Center Research Paper.

Dominitz, Jeff and Charles F Manski. 2007. "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study." *Journal of the European Economic Association* 5 (2-3):369–379.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2016. "Heterogeneity and Persistence in Returns to Wealth." NBER Working Paper.

Fagereng, Andreas, Magne Mogstad, Marte Rønning et al. 2015. "Why Do Wealthy Parents Have Wealthy Children." *Statistics Norway, Discussion Papers* 813.

Fletcher, Jason M. 2012. "Why Have Tobacco Control Policies Stalled? Using Genetic Moderation to Examine Policy Impacts." *PLoS ONE* 7 (12):1–6.

Fletcher, Jason M and Steven F Lehrer. 2011. "Genetic Lotteries within Families." *Journal of Health Economics* 30 (4):647–659.

Heckman, James J. 1995. "Lessons from the Bell Curve." *Journal of Political Economy* :1091–1120.

Heckman, James J and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2):145–149.

Hewitt, John K. 2012. "Editorial Policy on Candidate Gene Association and Candidate Gene-by-Environment Interaction Studies of Complex Traits." *Behavior Genetics* 42 (1):1–2.

Hudomiet, Peter, Gábor Kézdi, and Robert J Willis. 2011. "Stock Market Crash and Expectations of American Households." *Journal of Applied Econometrics* 26 (3):393–415.

Hurd, Michael D. 2009. "Subjective Probabilities in Household Surveys." *Annual Review of Economics* 1:543.

Jones, Charles I. 2015. "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality." *The Journal of Economic Perspectives* 29 (1):29–46.

Kahneman, Daniel and Amos Tversky. 1972. "Subjective Probability: A Judgment of Representativeness." In *The Concept of Probability in Psychological Experiments*. Springer, 25–48.

Kézdi, Gábor and Robert J Willis. 2003. "Who Becomes a Stockholder? Expectations, Subjective Uncertainty, and Asset Allocation." Unpublished manuscript, University of Michigan.

———. 2009. "Stock Market Expectations and Portfolio Choice of American Households." Unpublished manuscript, University of Michigan.

———. 2013. "Expectations, Aging and Cognitive Decline." In *Discoveries in the Economics of Aging*. University of Chicago Press, 305–337.

Kindermann, Fabian and Dirk Krueger. 2014. "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk." NBER Working Paper.

Lien, Diana S and William N Evans. 2005. "Estimating the Impact of Large Cigarette Tax Hikes: The Case of Maternal Smoking and Infant Birth Weight." *Journal of Human Resources* 40 (2):373–392.

Lillard, Lee and Robert J Willis. 2001. "Cognition and Wealth: The Importance of Probabilistic Thinking." *Michigan Retirement Research Center Working Paper* .

Locke, Adam E., Bratati Kahali, Sonja I. Berndt, Anne E. Justice, Tune H. Pers et al. 2015. "Genetic Studies of Body Mass Index Yield New Insights for Obesity Biology." *Nature* 518 (7538):197–206.

Okbay, Aysu, Jonathan P Beauchamp, Mark Alan Fontana, James J Lee, Tune H Pers, Cornelius A Rietveld, Patrick Turley, Guo-Bo Chen, Valur Emilsson, S Fleur W Meddens et al. 2016. "Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment." *Nature* 533 (7604):539–542.

Papageorge, Nicholas W and Kevin Thom. 2016. "Genes, Education and Labor Outcomes: Evidence from the Health and Retirement Study." IZA Discussion Paper 10200.

Poterba, James M and David A Wise. 1998. "Individual Financial Decisions in Retirement Saving Plans and the Provision of Resources for Retirement." In *Privatizing Social Security*. University of Chicago Press, 363–401.

Quadrini, Vincenzo. 2000. "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics* 3 (1):1–40.

Rietveld, Cornelius A, Sarah E Medland, Jaime Derringer, Jian Yang, Tõnu Esko, Nicolas W Martin, Harm-Jan Westra, Konstantin Shakhbazov, Abdel Abdellaoui, Arpana Agrawal et al. 2013. “GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment.” *Science* 340 (6139):1467–1471.

Rohwedder, Susann and Robert J Willis. 2010. “Mental Retirement.” *The Journal of Economic Perspectives* 24 (1):119–138.

Saez, Emmanuel and Gabriel Zucman. 2014. “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data.”

Savage, L.J. 1954. *The Foundations of Statistics*. Wiley.

Thompson, Owen. 2014. “Economic Background and Educational Attainment: The Role of Gene-Environment Interactions.” *Journal of Human Resources* 49 (2):263–294.

Thorgeirsson, Thorgeir E, Daniel F Gudbjartsson, Ida Surakka, Jacqueline M Vink, Najaf Amin, Frank Geller, Patrick Sulem, Thorunn Rafnar, Tõnu Esko, Stefan Walter et al. 2010. “Sequence Variants at CHRNB3-CHRNA6 and CYP2A6 Affect Smoking Behavior.” *Nature Genetics* 42 (5):448–453.

Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie. 2011. “Financial Literacy and Stock Market Participation.” *Journal of Financial Economics* 101 (2):449–472.

Vilhjalmsson, et al., Bjarni J. 2015. “Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores.” *The American Journal of Human Genetics* 87:576–592.

Vissing-Jørgensen, Annette. 2002. “Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures.” NBER Working Paper.

Willis, Robert J. and Sherwin Rosen. 1979. “Education and Self-Selection.” *Journal of Political Economy* 87 (5):S7–S36.

Wiswall, Matthew and Basit Zafar. 2015. “Determinants of College Major Choice: Identification Using an Information Experiment.” *Review of Economic Studies* 82 (2):791–824.

Yogo, Motohiro. 2016. “Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing, and Risky Assets.” *Journal of Monetary Economics* 80:17–34.

8 Tables and Figures

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	N
Male	0.42	0.49	4433
< 1930	0.32	0.47	4433
1930-1934	0.19	0.39	4433
1935-1939	0.21	0.41	4433
1940-1944	0.15	0.36	4433
1945-1949	0.07	0.26	4433
1950-1954	0.05	0.22	4433
Education (Years)	12.88	2.57	4433
None	0.16	0.37	4429
GED	0.05	0.21	4429
High School	0.53	0.5	4429
College (2 year)	0.04	0.21	4429
College (4 year)	0.13	0.33	4429
Masters	0.07	0.25	4429
Advanced	0.02	0.14	4429
Yrs. of ed: Father	9.76	3.61	3193
Yrs. of ed: Mother	10.26	3.11	3329

Notes: Summary statistics for our main cross-sectional sample of financial respondents.

Table 2: Wealth Distribution

	p10	p25	p50	p75	p90	Mean	St Dev
Wealth (Winz)	29,550	105,089	284,176	631,047	1,235,471	540,640	826,944
Wealth (No Housing)	12,861	45,000	146,686	413,621	919,036	411,180	1,355,300
Wealth (No Pensions)	1,212	66,144	224,520	549,822	1,124,892	495,831	1,083,248
Wealth (No H or P)	0	5,469	85,422	327,315	810,000	330,375	950,170

Notes: Wealth mean and distribution (10th, 25th, 50th, 75th and 90th percentiles) for total wealth, non-housing wealth, non-pension wealth and wealth that includes neither pensions nor housing. These statistics are calculated for the full sample of 15,517 household-year observations with non-missing wealth data.

Table 3: Components of Wealth

	p50	**p75**	**p90**	**Mean**	**Med Share**	**Mean Share**
Ret Plans (Employer)	0	0	0	22,343	0 %	4 %
Ret Inc (PV)	33,641	74,736	136,550	58,461	21 %	9 %
Real Estate	0	0	57,717	46,671	0 %	7 %
Business	0	0	0	36,742	0 %	6 %
IRAs	0	55,055	202,557	72,207	0 %	12 %
Stocks	0	28,000	214,708	89,883	0 %	14 %
Cash Equiv	8,102	30,000	81,023	33,169	5 %	5 %
CDs	0	5,408	60,000	22,980	0 %	4 %
Bonds	0	0	0	14,667	0 %	2 %
Other Assets	0	0	16,224	14,382	0 %	2 %
Other Debts	0	0	5,000	2,706	0 %	0 %
Trusts	0	0	0	2,380	0 %	0 %
Home Value	118,131	208,100	354,475	164,562	74 %	26 %
Mortgage	0	0	58,742	16,499	0 %	3 %
Home Loan	0	0	0	2,315	0 %	0 %
Second Home	0	0	27,704	21,964	0 %	4 %
Second Morgt.	0	0	0	1,450	0 %	0 %

Notes: Summary statistics of different sources of wealth (mean and distribution, including the 50th, 75th and 90th percentiles). Columns 5 and 6 are median and mean share, respectively, of each component in total wealth.

Table 4: The Polygenic Score and Wealth

Panel A: log Wealth			
	(1)	(2)	(3)
EA2Score	0.235***	0.103***	0.099***
	(0.022)	(0.022)	(0.022)
Resp Education	No	Yes	Yes
Parental Education	No	No	Yes
Obs.	15202	15202	15202
R^2	0.151	0.255	0.259
Panel B: Wealth (level)			
	(1)	(2)	(3)
EA2Score	117,301***	57,475***	54,520***
	(12,938)	(12,260)	(12,290)
Resp Education	No	Yes	Yes
Parental Education	No	No	Yes
Obs.	15517	15517	15517
R^2	0.089	0.169	0.172

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All regressions include the following *standard set of controls*: first ten principal components of the genetic data, a full set of birth year dummies, age dummies, calendar year dummies, a male dummy, interactions between the birth year and male dummies, interactions between the age and male dummies, a dummy variable for individuals in 2002 with dormant retirement accounts, and an interaction between the EA Score and the indicator for dormant accounts. Column (2) adds controls for the financial respondent's own education: years of education, and a full set of dummies for degrees. Column (3) adds controls for parental education: years of education for the respondent's father and mother, respectively, along with dummy variables indicating missing values for either. Standard errors are clustered at the household level. We use data on all household-year observations where no individual in the household is working for pay and not retired

Table 5: Total Wealth, Income Flows, and Financial Decisions

Dep. Var: log Tot. Wealth	(1)	(2)	(3)	(4)	(5)
EA Score	0.120***	0.121***	0.106***	0.097***	0.061***
	(0.027)	(0.026)	(0.026)	(0.026)	(0.023)
Avg log HH Inc		0.347***	0.333***	0.349***	0.273***
		(0.038)	(0.038)	(0.039)	(0.033)
log Sum Inher.			0.022***	0.022**	0.014*
			(0.008)	(0.008)	(0.007)
Ever Rec Inher.			0.156	0.151	0.108
			(0.095)	(0.095)	(0.081)
Ever Own Bus.				0.303***	0.257***
				(0.054)	(0.046)
Owns Stocks					1.032***
					(0.041)
Obs.	6943	6943	6943	6943	6943
R^2	0.291	0.330	0.345	0.354	0.455

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The dependent variable in all specifications is the log of total wealth, as used in Table 4. All regressions include the standard set of controls outlined in the Notes to 4, as well as controls for the respondent's education (years of schooling and a fully set of degree dummies), and controls for mother's and father's education. Standard errors are clustered at the household level. To allow for comparability of coefficients across specifications, we restrict the sample in all specifications here to households with non-missing observations on average household income, inheritances, business ownership, and stock ownership. We use data on all household-year observations where no individual in the household is un-retired and working for pay.

Table 6: Polygenic Score and Stock Ownership

Dep. Var: Owns Stocks	(1)	(2)	(3)
EA2Score	0.037***	0.027***	0.028***
	(0.009)	(0.010)	(0.010)
Lag of log Wealth		0.156***	0.152***
		(0.008)	(0.009)
Avg log HH Inc			0.025**
			(0.012)
Obs.	7893	4938	4938
R^2	0.221	0.374	0.376

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The dependent variable in all specifications is a dummy variable indicating whether the household owns any stocks or stock mutual funds. All regressions include the standard set of controls outlined in the Notes to 4, as well as controls for the respondent's education (years of schooling and a fully set of degree dummies), and controls for mother's and father's education. Standard errors are clustered at the household level. We use data on all household-year observations where no individual in the household is un-retired and working for pay. For Columns (2)-(3), we restrict the sample to households with no working, un-retired individuals in the previous period so that lagged wealth is accurately measured.

Table 7: Risk Aversion, Wealth, and Stock Ownership

	(1)	(2)	(3)	(4)	(5)
Dep Var:	**Risk Averse**	**Tot. Wealth**	**Tot. Wealth**	**Owns Stocks**	**Owns Stocks**
EA Score	-0.013*	0.124***	0.121***	0.035***	0.034***
	(0.006)	(0.030)	(0.030)	(0.010)	(0.010)
Risk Averse			-0.167***		-0.073***
			(0.063)		(0.020)
Obs.	5321	7050	7050	7458	7458
R^2	0.102	0.270	0.272	0.216	0.220

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The dependent variable in Column (1) is a binary measure for risk aversion described in Section 4.3. The dependent variable in Column (2) is the log of total household wealth. The dependent variable in Columns (3)-(4) is a binary for any stock ownership. Since the risk aversion measure is time-invariant (whether an individual ever reported the most risk averse response), we only use one observation per individual and include a slightly different control set that includes the genetic principal components, birth year dummies, a male dummy, interactions between birth year dummies and the male dummy, and the own and parental education controls. All other regressions include the standard set of controls outlined in the Notes to 4, as well as controls for own and parental education. Standard errors are clustered at the household level. Note as well that the sample for Column (1) includes all individuals with non-missing risk aversion data, regardless of whether they are a financial respondent.

Table 8: Non-Financial Respondent Score and Household Wealth

Panel A: Spouse's Score				
	(1)	(2)	(3)	(4)
Dep Var:	**Tot. Wealth**	**Tot. Wealth**	**Owns Stocks**	**Owns Stocks**
EA Score	0.097***	0.092***	0.022*	0.019*
	(0.028)	(0.028)	(0.011)	(0.011)
EA Spouse		0.078***		0.034***
		(0.026)		(0.011)
Obs.	5166	5166	5182	5182
R^2	0.314	0.318	0.228	0.232
Panel B: Max Score				
	(1)	(2)	(3)	(4)
Dep Var:	**Tot. Wealth**	**Tot. Wealth**	**Owns Stocks**	**Owns Stocks**
EA Score	0.097***		0.022*	
	(0.028)		(0.011)	
Max EA Score		0.136***		0.042***
		(0.032)		(0.014)
Obs.	5166	5166	5182	5182
R^2	0.314	0.316	0.228	0.231

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The dependent variable in all specifications is the log of total household wealth. All other regressions include the standard set of controls outlined in the Notes to 4, as well as controls for own and parental education. Standard errors have been clustered at the household level. The samples for Columns (1) and (3) have been restricted to observations of financial respondents with non-missing values for the spousal EA Score.

Table 9: *EA Score* and Financial Literacy

Dep Var:	(1) **Compound Interest**	(2) **Real Interest**	(3) **Diversify**	(4) **All Correct (1)-(3)**	(5) **Inflation and Lending**
EA2Score	-0.001	0.032**	0.027	0.040*	0.080***
	(0.018)	(0.016)	(0.021)	(0.021)	(0.020)
Obs.	666	667	667	663	674
R^2	0.245	0.203	0.239	0.294	0.270

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. The dependent variables in Columns (1)-(3) are dummy variables indicating correct responses for the three questions that were included together in a financial literacy module in the 2010 wave of the HRS. The dependent variable in Column (4) aggregates these items by constructing a binary indicating whether or not individuals got all three questions correct. The dependent variable in Column (5) indicates a correct response to a separate module question (with different respondents) on inflation and lending (see text for details). All regressions include the standard set of controls outlined in the Notes to Table 4 (except for the dormant pension controls), as well as controls for own and parental education. Standard errors are clustered at the household level.

Table 10: *EA Score* and Beliefs

	(1) **Dev. from Objective**	(2) **0% Prob**	(3) **50% Prob**	(4) **100% Prob**
Panel A: Market Up				
EA2Score	-0.453***	-0.004***	-0.002	-0.003**
	(0.141)	(0.001)	(0.003)	(0.001)
Obs.	39626	39626	39626	39626
R^2	0.071	0.031	0.011	0.021
Panel B: U.S. Depression				
EA2Score	-0.343***	-0.005***	-0.002	-0.004**
	(0.125)	(0.002)	(0.003)	(0.002)
Obs.	32971	32971	32971	32971
R^2	0.064	0.030	0.022	0.044
Panel C: Double Digit Inf				
EA2Score	-0.612***	-0.007***	0.000	-0.007***
	(0.184)	(0.002)	(0.004)	(0.002)
Obs.	19541	19541	19541	19541
R^2	0.050	0.027	0.023	0.040

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All regressions include the standard set of controls outlined in the Notes to 4 (excluding the dormant pension controls), as well as controls for own and parental education. The samples for all regressions in this table include person-year observations on all respondents (not just financial respondents with non-missing wealth data). Standard errors are clustered at the household level.

Table 11: Beliefs and Household Wealth

	(1) **log Wealth**	(2) **log Wealth**	(3) **Owns Stocks**	(4) **Owns Stocks**
Ever Pr Mrkt Up 0%	-0.221***	-0.125**	-0.101***	-0.059***
	(0.055)	(0.064)	(0.016)	(0.020)
Ever Pr Mrkt Up 100%	0.257***	0.257***	0.099***	0.097***
	(0.064)	(0.063)	(0.021)	(0.021)
Ever Pr Rec 0%	0.006	0.038	-0.019	-0.012
	(0.055)	(0.056)	(0.017)	(0.018)
Ever Pr Rec 100%	-0.253***	0.001	-0.058***	0.002
	(0.059)	(0.079)	(0.018)	(0.025)
Ever DD Inf 0%	-0.084	-0.088	-0.000	-0.000
	(0.067)	(0.067)	(0.021)	(0.021)
Ever DD Inf 100%	-0.208***	-0.134	-0.092***	-0.069**
	(0.068)	(0.092)	(0.020)	(0.028)
Max Dev Mrkt. Up		-0.003**		-0.002***
		(0.001)		(0.000)
Max Dev Rec.		-0.009***		-0.002***
		(0.002)		(0.001)
Max Dev DD Inf		-0.001		-0.000
		(0.002)		(0.001)
Obs.	14045	14045	14510	14510
R^2	0.266	0.271	0.201	0.205

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All regressions include the standard set of controls outlined in the Notes to Table 4 (excluding the dormant pension controls), as well as controls for own and parental education. Standard errors are clustered at the household level.

Table 12: Household Wealth, Expectations, and Other Scores

Panel A: Correlations between Polygenic Scores:

	EA Score	BMI Score	Cigs. Score
EA Score	1.00		
BMI Score	-0.18	1.00	
Cigs. Score	-0.09	0.03	1.00

Panel B: Other Scores, Wealth, and Expectations:

	(1) **log Wealth**	(2) **Dev. from Objective: Market Up**	(3) **Dev. from Objective: Depression**	(4) **Dev. from Objective: Double Digit Inf.**
EA Score	0.086***	-0.451***	-0.312**	-0.580***
	(0.022)	(0.144)	(0.128)	(0.187)
BMI Score	-0.060***	0.080	0.140	0.015
	(0.021)	(0.140)	(0.120)	(0.183)
Cigs Score	-0.053**	0.070	0.132	0.047
	(0.021)	(0.138)	(0.120)	(0.177)
Obs.	15003	39171	32609	19325
R^2	0.260	0.071	0.065	0.050

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Panel A shows the cross-sectional correlations for all 8,459 individuals with non-missing values of the polygenic scores. Column (1) in panel B replicates the basic log wealth regression from Table 4 but now includes all three polygenic scores. Columns (2)-(4) in panel B replicate the specifications from Column (1) of Table 10, including all three polygenic scores.

Table 13: Pensions and Household Wealth

	(1) Has Pension	(2) log Wealth	(3) log Wealth	(4) log Wealth
EA Score	-0.004	0.095***	0.137***	0.117***
	(0.008)	(0.022)	(0.034)	(0.034)
DB Pension		0.343***	2.401**	1.701
		(0.038)	(1.100)	(1.083)
EA Score x DB			-0.079**	-0.059
			(0.038)	(0.038)
Max Dev Mrkt. Up				-0.007***
				(0.002)
Max Dev Rec.				-0.011***
				(0.002)
Max Dev DD Inf				-0.005***
				(0.002)
(Max Dev Mrkt. Up) x DB				0.004*
				(0.002)
(Max Dev Rec.) x DB				0.005**
				(0.003)
(Max Dev DD Inf) x DB				0.004**
				(0.002)
Obs.	14045	14045	14045	14045
R^2	0.084	0.266	0.271	0.289

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All specifications start with the basic specification for log wealth in Column (3) of Table 4. Standard errors are clustered at the household level. Columns (2)-(4) also include a full set of interactions between the principal components of the genetic data and the defined benefit pension dummy. This explains the large coefficient on DB Pension in those specifications.

Table 14: Beliefs and Childhood SES

	(1) **Not Poverty**	(2) **Income**	(3) **Move or Help**	(4) **Father Unemployed**
Panel A: Stock Market Up				
High SES	-13.868*	3.028	16.837**	5.196
	(7.621)	(8.036)	(8.059)	(7.769)
EA Score	-0.779***	-0.727***	-0.778***	-0.714**
	(0.282)	(0.219)	(0.284)	(0.280)
EA $\times$ High SES	-0.305	-0.381	-0.352	-0.387
	(0.323)	(0.320)	(0.324)	(0.321)
Obs.	39035	30533	38987	39145
R^2	0.059	0.057	0.059	0.057
Panel B: Recession				
High SES	-4.048	-0.810	8.171	-0.081
	(6.594)	(6.868)	(7.137)	(6.854)
EA Score	-0.763***	-0.556***	-0.720***	-0.599**
	(0.237)	(0.192)	(0.245)	(0.243)
EA $\times$ High SES	0.179	-0.092	0.134	-0.031
	(0.277)	(0.273)	(0.282)	(0.282)
Obs.	32509	26164	32419	32569
R^2	0.060	0.060	0.061	0.059
Panel C: Double Digit Inflation				
High SES	-12.452	15.732	-3.961	-1.200
	(9.971)	(9.721)	(11.293)	(10.121)
EA Score	-0.636*	-0.511**	-0.936**	-0.759**
	(0.361)	(0.259)	(0.382)	(0.357)
EA $\times$ High SES	-0.508	-1.215***	-0.136	-0.391
	(0.416)	(0.375)	(0.431)	(0.413)
Obs.	19290	17454	19254	19312
R^2	0.043	0.045	0.043	0.042

Notes: Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. All specifications start with the basic specification for deviations from objective beliefs found in Column (1) of Table 10. Standard errors are clustered at the household level. All specifications include a full set of interactions between the principal components of the genetic data and the High SES dummy. This explains the large coefficients on High SES dummy in those specifications.

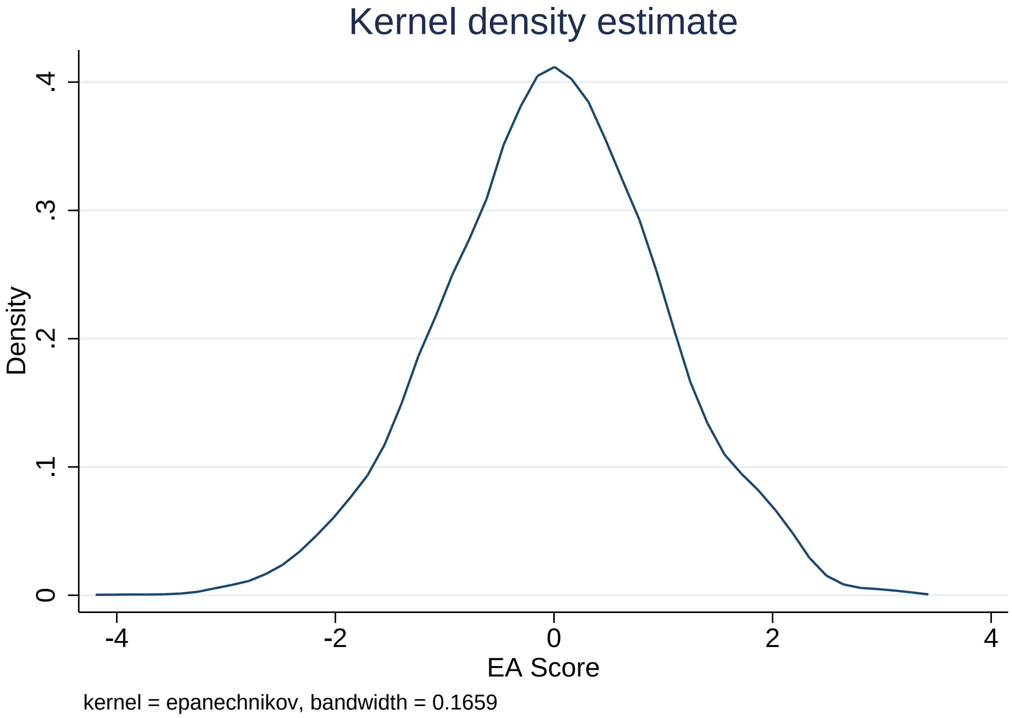

Figure 1: *Notes*: EA Score Distribution among HRS Individuals.

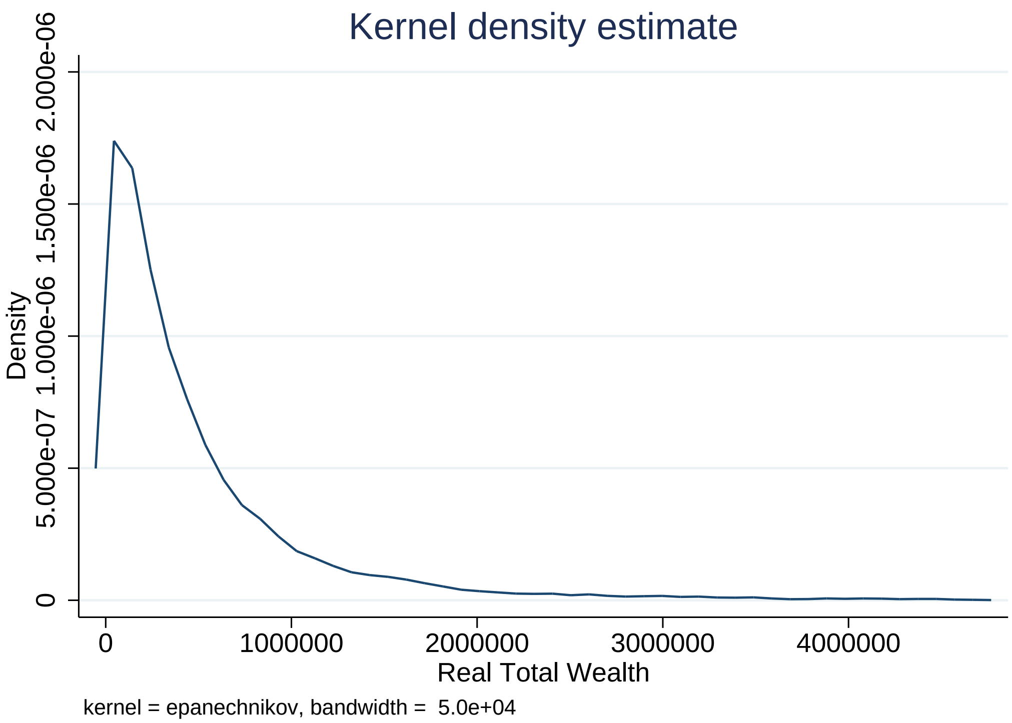

Figure 2: *Notes*: Wealth Distribution among HRS Individuals.

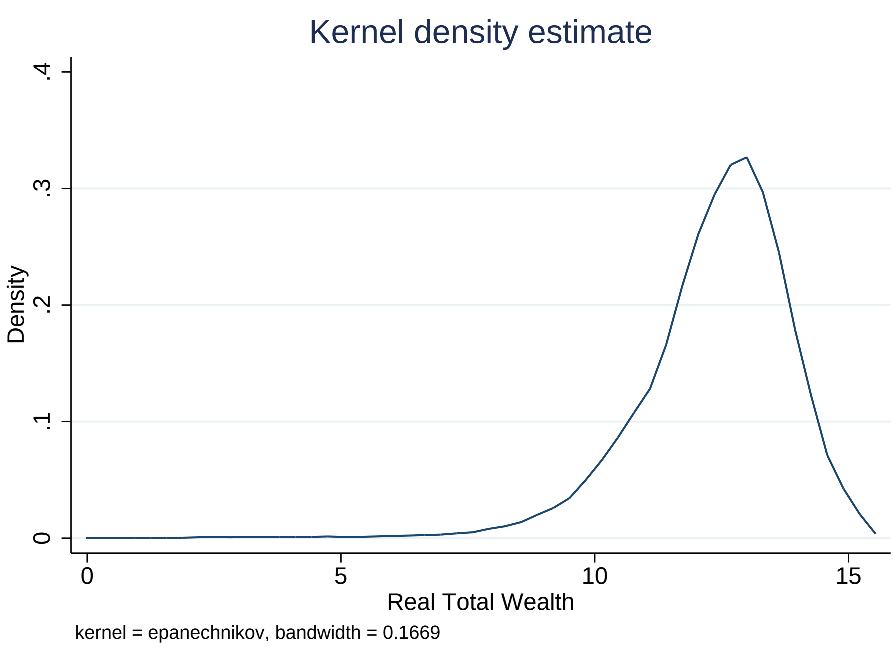

Figure 3: *Notes*: Log Wealth Distribution among HRS Individuals.

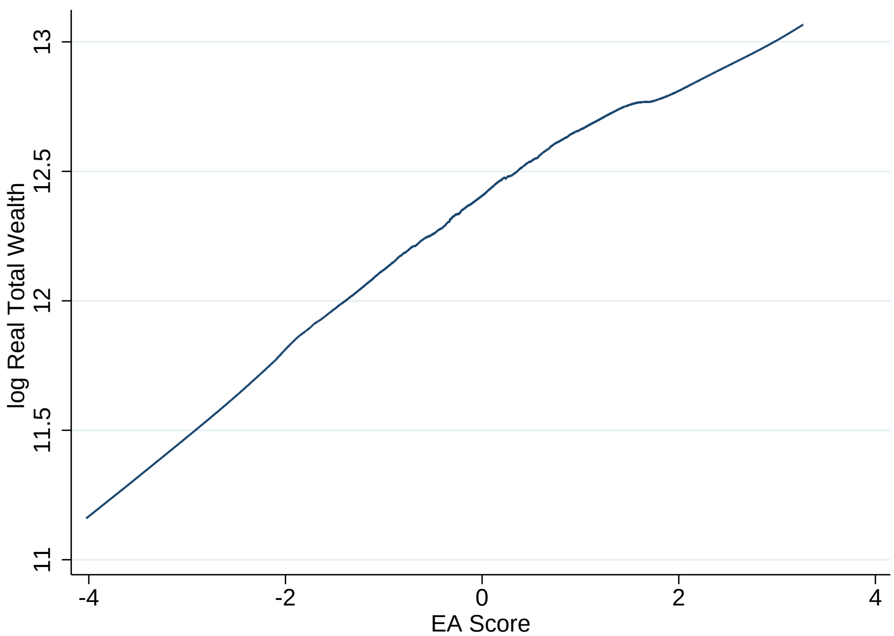

Figure 4: *Notes*: EA Score and Wealth among HRS Individuals.

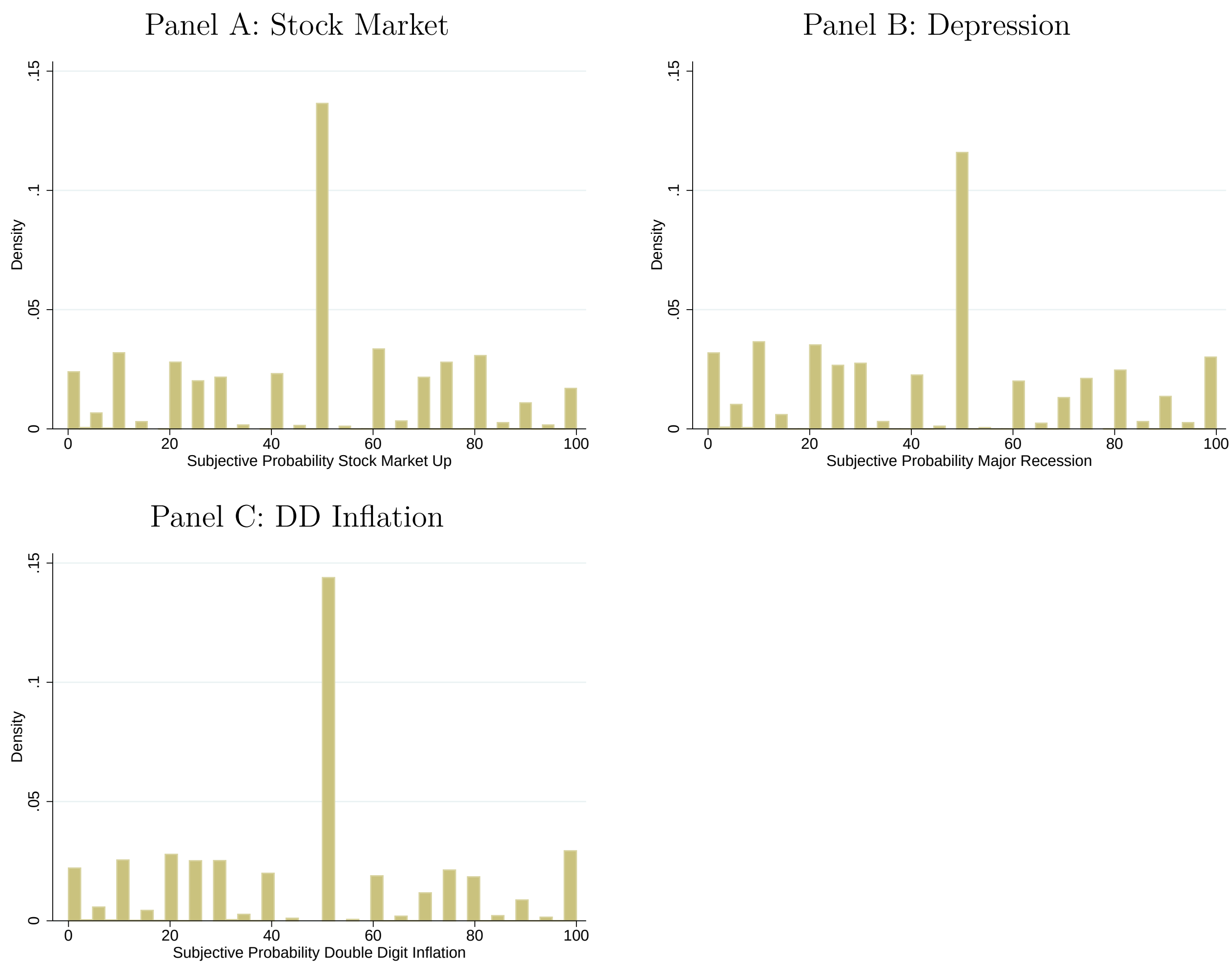

Figure 5: *Notes*: Distributions of Subjective Beliefs.

Online Appendix:

"Genetic Ability, Wealth, and Financial Decision-Making"

By: Danny Barth, Nicholas W. Papageorge and Kevin Thom

A Additional Details on GWAS and Construction of the EA Score

What follows is nearly identical to the genetic data appendix in an earlier paper, Papageorge and Thom (2016), which uses the same genetic score used in the current paper. We reprint the appendix here solely for the reader's convenience In this appendix, we provide a brief introduction to molecular genetics and the kinds of genetic data that we use in this study. We repeat some portions of Section 2 so that this appendix can provide a self-contained introduction to GWAS and the EA Score used in our analysis. First, we describe some basic features of the human genome. Next, we discuss how statistical gene-discovery projects can produce scores that are useful for the prediction of economic outcomes such as educational attainment. We highlight how recent advances permit credible and replicable inference.

The human genome consists of approximately 3 billion nucleotide base pairs spread out over 23 chromosomes. Each individual possesses two copies of each chromosome, one from each parent. A gene is a subsequence of base pairs within a chromosome. On average, each gene is made up of over 100,000 base pairs. Each base pair can either be an adenine-thymine (AT) pair, or a guanine-cytosine (GC) pair. Thus, the human genome can be thought of as a series of 3 billion genetic addresses, each of which can contain one of two nucleotide pairs.

A particular location in the genome can be referred to by a name (e.g. rs7937), which indicates its position in the genome. At the vast majority of such locations (about 99%), there is no variation in the observed nucleotide pair across humans or across chromosomes within a human. A single-nucleotide polymorphism (SNP) exists when there are differences in the nucleotide pair present at a particular location on the genome. An allele refers to one of the variants that may be present at a particular SNP. If AT is more commonly found at a particular SNP, it is referred to as the major allele, and then GC is referred to as the minor allele.

A traditional approach to the discovery of gene-behavior associations rests on examining *candidate genes*. Under this paradigm, researchers use some knowledge of the relevant biological processes to suggest places in the genome that might contain SNPs associated with a particular outcome. Unfortunately, this approach to identifying gene-economic outcomes has also generated a large number of reported associations that have failed to replicate outside of their discovery samples. This problem has been so widespread that an editorial statement from the journal *Behavior Genetics* stated that "The literature on candidate gene associations is full of reports that have not stood up to rigorous replication," and that "it now seems likely that many of the published findings of the last decade are wrong or misleading and have not contributed to real advances in knowledge," (Hewitt, 2012). This pattern has emerged, in part, because traditional candidate gene studies have been severely underpowered to detect real genetic effects. Sample sizes in general have been too small relative to the true effect sizes of individual SNPs, making it likely that statistically significant associations are the result of chance. This problem is exacerbated when studies search over many candidate genes, creating a multiple hypothesis testing problem that increases the likelihood of finding false positive results (Benjamin et al., 2012).

An alternative to candidate genes is an approach called a genome-wide association study (GWAS). Under the GWAS methodology, researchers scan the entire genome for SNPs that are associated with a particular phenotype (trait or outcome), but adopt strong measures to deal with multiple hypothesis-testing. For a particular outcome of interest, y_i, and for a set of observed SNPs, $\{SNP_{ij}\}_{j=1}^{N^J}$, a GWAS study proceeds by obtaining estimates of N^J separate regressions of the form:

$$y_i = \mu X_i' + \beta_j SNP_{ij} + \epsilon_{ij} \tag{1}$$

Here SNP_{ij} measures the number of copies of a reference allele possessed by individual i for SNP j. For example, if the reference allele at SNP j is AT, then SNP_{ij} could take the values 0, 1, or 2. The maximum value of 2 reflects the fact that an individual can have at most two copies of the reference allele — one on each inherited chromatid. Additionally, X_i is a vector of controls, including principal components of the genetic variables $\{SNP_{ij}\}_{j=1}^{N^J}$. Principal components of the genetic data are added to control for population stratification. For example, it could be that SNP_{ij} is correlated with a particular ethnicity or ancestry group. Failure to control for the principal components could generate observed SNP-phenotype relationships that reflect the influence of broader ethnic differences rather than the influence of a particular genetic marker.

After obtaining estimates for all N^J versions of equation (1), those estimated coefficients $\widehat{\beta}_j$ with sufficiently small p-values are said to reflect relationships that are genome-wide significant. Given the huge number of regressions run under this methodology, the significance thresholds in modern GWAS are typically very strict. A conventional threshold is 5×10^{-8}. This approach has become popular and as a consequence of its stringency requirements, has led to the discovery of a number of credible genetic associations. For example, the well-known FTO gene for obesity was discovered through a GWAS, despite the lack of any existing biology that would have suggested it as a candidate gene (Benjamin et al., 2012).

Existing work has demonstrated the importance of credibly identified SNPs for several economic outcomes. These SNPs either directly emerged from a GWAS, or were candidate genes that were validated by later GWAS results. An established literature documents a number of credible genetic associations with smoking behaviors (Bierut, 2010; Thorgeirsson et al., 2010). Fletcher (2012) demonstrates that a SNP associated with smoking intensity also appears to moderate the effect of tobacco taxes. More closely related to our work, another set of studies suggests indirect linkages between genetic variants and human capital. For example, Fletcher and Lehrer (2011) use a set of SNPs associated with health outcomes to provide exogenous within-family variation to estimate a causal relationship between health and education. Finally, Thompson (2014) shows that a variant associated with the MAOA gene appears to moderate the relationship between income and education.

Recent work using GWAS has discovered some of the first direct associations between specific SNPs and education. Rietveld et al. (2013) identified three SNPs (rs9320913, rs11584700, rs4851266) attaining genome-wide significance in a GWAS for educational attainment. Follow-up work by the same team (the Social Science and Genetics Consortium) has recently extended the Rietveld et al. (2013) study to perform an educational attainment GWAS with a sample size of 293,723. This follow-up study, Okbay et al. (2016), has discovered 74 SNPs that attain genome-wide significance. We build our analysis here on the gene-education associations found in this follow-up study.

One common technique adopted in the GWAS literature is to take observed SNPs and the estimated GWAS coefficients (the $\widehat{\beta}_j$) and aggregate them into a polygenic score that can be used for prediction. Typically these scores take the following form:

$$PGS_i = \sum_j \widetilde{\beta}_j SNP_{ij} \tag{2}$$

where $\widetilde{\beta}_j$ is some transformation of the underlying GWAS coefficients. The $\widehat{\beta}_j$ estimates are typically corrected to account for correlation between SNPs and prevent over or under prediction. The follow up study Okbay et al. (2016) combines all genotyped SNPs into a polygenic score that attains a predictive power of up to 3.85% of the variation in educational attainment.[29] In our study, we use SNP weights $\widehat{\beta}_j$ that have been adjusted using a technique called LD Pred (Vilhjalmsson, 2015), and applied to the genetic data in the HRS.[30] We refer

[29] We note as well that the polygenic score that we use in this study combines all SNPs analyzed in Okbay et al. (2016), not just those reaching genome-wide significance. As noted in Okbay et al. (2016), this maximizes the predictive power out of sample.

[30] We would like to especially thank Aysu Okbay, a member of the Social Science and Genetics Consortium,

to the polygenic score created using these weights as the *EA Score*, where "EA" stands for "educational attainment". We refer to it this way since other polygenic scores exist which capture genetic variation explaining different outcomes.

B Data Issues

This appendix provides details on the construction of our wealth data and our measurement of stock market participation. Our data are largely constructed from the RAND wealth and income files. The RAND files are carefully cleaned and consistently coded by RAND Corportation and are available for public use. The RAND files have been used in both academic and industry publications, and ensure comparability and consistency across HRS waves and research projects. We refer the reader to the RAND codebook and documentation for further details.

One important shortcoming of the RAND wealth files is the exclusion of employer-sponsored retirement plan account balances. While the RAND wealth files do include the balances of IRAs and other non-employer-sponsored plans, wealth accumulated in employer-sponsored 401k, 403(b), and other such accounts are not included. For households at or near retirement, such accounts can be a significant source of wealth. Further, such accounts may be the only vehicles through which households invest in the stock market, and measures of stock market participation will understate true participation if these plans are not considered.

Unfortunately, data on employer-sponsored retirement plans are not asked in every wave, and are sometimes inconsistently coded across waves. The remainder of this section focuses on our methodology for coding retirement account balances and stock market participation inferred from those accounts.

B.1 Wealth in Retirement Accounts

Broadly speaking, there are two types of retirement plans: defined benefit plans, such as traditional pensions (which the HRS calls type A plans), and defined contribution plans, such as 401k and 403(b) plans (which the HRS calls type B plans). We discuss each type of plan in turn.

for graciously generating and sharing this score with us.

B.1.1 Defined Benefit Plans

To deal with issues arising from type A style retirement plans, our sample includes only households fully in retirement (households in which no member of the household is currently working). We exclude working households because expected benefits from defined-benefit pension plans are likely to be both an important source of wealth and noisily measured. For retired households, our assumption is that those who report receiving pension income were included in defined-benefit pension plans at some point during their working lives, and those who do not receive pension income in retirement were not included in such plans. To the extent that households misreport pension income, for example if income from an annuity converted from a 401k plan is reported as pension income, or if households have delayed receiving pension benefits until some future date, our assignment of households participating in type A plans will be biased. Further, because the household earns a guaranteed stream of income regardless of the underlying investments that support that income (and because we do not observe these underlying investments), we do not consider a household's participation in type A pension plans to be participation in the stock market.

We include retirement income in our household wealth measure by calculating the price of an actuarially fair annuity based on the entirety of household retirement income, which includes pension income, annuity income, and income from social security. We follow Yogo (2016) by calculating the present discounted value of this income based on a 1.5% annual risk-free rate of return, and discount income in each year by the probability of the recipient surviving until that year.[31] Specifically, we calculate the present value of retirement income, P_t, as:

$$P_t = Y_t \sum_{s=1}^{T-t-1} \frac{\prod_{u=1}^{s} p_{t+u}}{R}, \tag{3}$$

where Y_t is total retirement income, p_t is the recipient's survival probability in period t and is a function of gender, birth cohort, and age, and $R = 1.015$ is the annual risk-free rate of return.

B.1.2 Defined Contribution Plans

Wealth in defined-contribution style plans is a bit trickier. Households may have plans associated with multiple previous employers. To calculate comprehensive measures of wealth and stock market participation, we would like to know both the balances and asset allocations

[31] We differ from Yogo (2016) in that we use the probability of death of the individual receiving the income, rather than of the female spouse.

of all employer-sponsored type B plans from all previous jobs. Unfortunately, this is not always possible.

In years 1996, 1998, and 2002-2010 (comprising even-numbered years), we have the highest quality data on total balances in employer-sponsored type B retirement plans.[32] In these years, our wealth data include balances of employer-sponsored plans that are still maintained through that employer, and have not been converted to annuities or rolled over into IRAs. The HRS refers to such plans as *dormant plans*. Unfortunately, the value of dormant plans at employers prior to retirement are not asked in 1992, 1994, and 2000.

Dormant plans also present problems for measurement of stock market participation. While in years 2002-2010 the stock allocation within a respondent's retirement plan at the current employer is observable for working households, the stock allocation in dormant plans for retired households is not. This means our stock market participation variable does not include stock ownership in dormant plans. The stock market participation variable is determined only by information in the assets and income section of the data, which comprises only stock and stock mutual funds as well as the stock allocation in IRA and Keogh accounts.

[32] In 2012, the pension data were changed to an entirely new format.