Molecular Genetics, Risk Aversion, Return Perceptions, and Stock Market Participation

March 13, 2019

Richard Sias, Laura Starks, and H. J. Turtle*

ABSTRACT

Molecular genetic endowments related to cognition, personality, health, and body shape, established at least half a century prior, predict an individual's risk aversion, beliefs regarding the distribution of expected equity returns, and equity market participation. We estimate that approximately 30% of the relation between equity market participation and both risk aversion and beliefs regarding the distribution of expected equity returns arises from the portion of risk aversion and beliefs associated with these genetic endowments. Molecular genetic endowments are also strongly associated with many of the variables (e.g., trust, sociability, wealth) shown to explain heterogeneity in equity market participation.

* Sias is from the Department of Finance, Eller College of Management, University of Arizona, Tucson, Arizona, 85721; (520) 621-3462, sias@eller.arizona.edu. Starks is from the Department of Finance, McCombs School of Business, University of Texas at Austin, Austin, Texas 78712, (512) 471-5899, lstarks@mail.utexas.edu. Turtle is from the Department of Finance and Real Estate, College of Business, Colorado State University, Fort Collins, Colorado, 80523, (970) 491-3596, harry.turtle@colostate.edu.

Molecular Genetics, Risk Aversion, Return Perceptions, and Stock Market Participation

"*Oh there ain't no other way, Baby I was born this way*"

-Lady Gaga

Traditional utility maximizing theory concludes nearly all investors should invest some portion of their wealth in equity markets (e.g., Merton, (1969, 1971)) and most should hold almost all savings in equity (e.g., Heaton and Lucas (1997)). Contrary to this work, however, individuals exhibit substantial heterogeneity in portfolio allocations (see, for example, Curcuru, Heaton, Lucas, and Moore (2010)) and, particularly puzzling, many individuals hold no equities. Broadly, most research suggests two potential explanations for observed heterogeneity in equity participation. First, according to traditional economic theory, the fraction of assets allocated to equities is a function of an investor's risk aversion and beliefs about the distribution of expected returns. Thus, heterogeneity in risk aversion or return beliefs can help explain the variation in their choices. Second, factors outside of traditional economic models—specifically, heterogeneity in circumstances (e.g., participation costs, background risks) may also help explain the observed participation heterogeneity. As a result, a large literature has evolved demonstrating that variation in a number of individual characteristics—such as wealth, trust, sociability, life experience, and even body size and shape—can help explain the heterogeneity in equity market participation either because these variables help explain differences across factors driving traditional theory (i.e., risk aversion and beliefs) or because these variables help explain differences in circumstances (e.g., participation costs).

However, underlying the individual characteristics that can explain heterogeneity in portfolio decisions are genetic factors.[1] For example, researchers have inferred the genetic component of portfolio decisions by employing the classic "twins" approach, that is, by comparing the similarity of monozygotic (i.e., 'identical') twins to dizygotic (i.e., 'fraternal') twins in the Swedish Twins Registry. Further, recent advances in molecular genetics has allowed scientists to identify genetic markers associated with specific phenotypes (i.e., observable characteristics or traits).[2] This work has led genetic researchers to move beyond twin studies and develop a deeper understanding regarding the biological origin and nature of behaviors and choices—the specific genetic markers associated with a phenotype. In this study, we further develop our understanding the impact of genetic endowments on financial decisions. We examine whether specific molecular genetic factors relate to traditional

[1] Harlow and Brown (1990) provide some of the earliest evidence linking individual risk tolerance to underlying genetic/biological characteristics.

[2] The most relevant advances for our purposes are low cost genotyping and the development of consortium-based large scale Genome Wide Association Studies (GWAS).

economic theory through heterogeneity in risk aversion, beliefs regarding the distribution of equity returns, and ultimately, in observed equity market participation. Using a large panel data set from the Health and Retirement Study that includes individual financial, psychological, social, and biological data (i.e., DNA characteristics), our analysis allows us to not only attribute heterogeneity in risk aversion, return beliefs, and equity market participation to specific genetic factors, but also to estimate the role of those genetic endowments in linking equity market participation to risk aversion, beliefs, or variables known to explain equity market participation (e.g., trust, sociability).

We restrict analysis to eight genetic endowments that are related to cognition, personality, health, and body shape, based on previous work linking observed characteristics to equity market participation. Specifically, we examine composite genetic scores (known as polygenic or genome-wide scores) associated with cognition (Educational Attainment, General Cognition), personality (Neuroticism, Depressive Symptoms), health (Myocardial Infarction, Coronary Artery Disease), and body shape (BMI, Height).[3]

We begin our analysis by focusing on the role of these eight composite genetic scores in explaining heterogeneity in risk aversion and return distribution beliefs, i.e., the factors driving heterogeneity in equity participation within the traditional utility maximizing framework. Because the genome is determined at conception, and our sample is limited to individuals aged 50 to 80, these genetic scores are established at least a half century prior to our observations of individual risk aversion or beliefs.[4]

We find that these genetic endowments predict observed heterogeneity in traditional economic model inputs. Specifically, an individual's self-rated risk aversion is inversely related to their genetic scores associated with cognition and body shape (Educational Attainment and Height) and positively related to their genetic scores associated with personality (Neuroticism and Depressive Symptoms) and poor health (Myocardial Infarction and Coronary Artery Disease).[5] Moreover, we find that individual's genetic endowments are not only associated with their expressed risk aversion, but the endowments also predict heterogeneity in individual beliefs regarding the distribution of expected

[3] Because we examine both genotype (information from the genome) and phenotype (observed traits), we capitalize genotype variables to reduce confusion, e.g., "General Cognition" refers to the genetic endowment associated with cognition, often measured by a polygenic score (PGS), where "cognition" refers to an estimate of an individual's observed cognitive performance.

[4] Although a link between economic outcomes and genetic endowments may seem very temporally disjoint, work suggests genetic impacts may span millennia (cf., Ashraf and Galor (2013)).

[5] Several of the genetic scores we examine are strongly correlated (e.g., Coronary Artery Disease and Myocardial Infarction). To facilitate a connection with the extant literature, we begin our discussion with the observed simple relations between outcomes (risk aversion, beliefs, and equity market participation) and each of the genetic endowments individually. Our latter analysis also considers the joint, or simultaneous, impact of all genetic scores on outcomes.

equity returns. Specifically, beliefs regarding the likelihood the stock market will increase over the next year are positively related to genetic scores associated with cognition (Educational Attainment and General Cognition) and negatively related to genetic scores associated with personality and health (Neuroticism and Myocardial Infarction). Similarly, beliefs regarding the likelihood of a 20% or greater decline in the market over the next year (i.e., a measure of a given individual's perception of the riskiness of the expected return distribution) are negatively related to genetic scores associated with cognition (Educational Attainment and General Cognition), but positively related to genetic scores associated with personality (Neuroticism and Depressive Symptoms).[6]

These significant relationships between genetic endowments and both risk aversion and return beliefs suggest that these endowments should also be related to individual's choices regarding equity market participation. Not surprisingly, we find that the same genetic endowments also predict equity market participation—individuals with higher genetic endowments associated with cognition and body shape (Educational Attainment, General Cognition, and Height) are more likely to invest in equity markets (and in addition invest a larger fraction of their wealth in risky assets). In contrast, individuals with higher genetic scores associated with personality (Neuroticism and Depressive Symptoms), health (Myocardial Infarction and Coronary Artery Disease), and body shape (BMI) are less likely to participate in equity markets (and invest a lower fraction of their wealth in risky assets). We caution that, analogous to evidence linking genetic scores to educational attainment (e.g., Okbay et al. (2016)), we do not suggest that there exists an "equity market participation gene" any more than there exists a gene that causes individuals to graduate from college. Rather, we propose that individuals' genetic endowments correlate with equity market participation, at least in part, through genetic influences on the willingness to take risk, and on perceptions of the distribution of equity returns. These influences may be direct or indirect. For instance, the genetic endowments associated with Neuroticism may directly impact risk aversion or they may interact with environment to impact risk aversion. For instance, a genetic endowment for high Neuroticism may result in lower sociability and, as Hong, Kubik, and Stein (2004) point out, lower sociability may result in individuals being less likely to "learn" about the historically high returns offered by equity markets (i.e., gene by environment interactions).

[6] We treat beliefs regarding the likelihood of a greater than 20% fall in equity prices as a measure of an individual's view of the perceived riskiness of equity investing. We recognize, however, that such beliefs may also relate to risk aversion or ambiguity. For example, as Malmendier and Nagel (p. 381, footnote 6, 2011) point out, an investor with optimistic beliefs might also naturally report less risk aversion. Dow and Werlang (1992) provide a seminal examination of how ambiguity can impact portfolio decisions.

We seek to apportion attribution between genes and the environment through subsequent tests. Thus, our third set of tests estimates the role of the eight genetic factors in linking equity market participation to risk aversion or return beliefs. We begin by demonstrating, consistent with traditional economic theory, that equity market participation is: (1) inversely related to self-rated risk aversion, (2) positively related to beliefs regarding the likelihood of a positive market return over the next year, and (3) inversely related to beliefs regarding the likelihood of a greater than 20% decline in equity prices over the next year. We then partition risk aversion and return beliefs into two components, one predicted by the eight molecular genetic endowments and the other portion independent of the molecular genetic endowments. We examine the relation between equity market participation and both the genetic and non-genetic components of risk aversion and beliefs. We estimate that, when combined, 30% of the relation between equity market participation and risk aversion and return beliefs arises from the "genetic" portion of risk aversion and return beliefs.

We next examine whether there exists a genetic source for the previously identified characteristics that help explain the heterogeneity in equity market participation. To do so, we focus on 11 characteristics—wealth, income, education, cognition, trust, sociability, optimism, growing up poor, height, BMI, and health—that previous work identifies as helping explain heterogeneity in equity market participation. We differentiate these characteristics from risk aversion and beliefs because these explanatory variables are nearly uniformly hypothesized to impact heterogeneity in equity market participation via influencing heterogeneity in risk aversion, beliefs, or factors outside of classical models (i.e., heterogeneity in circumstances). For instance, Malmendier and Nagel (2011) hypothesize that an individual's life experience (e.g., growing up in a depression) impacts an individual's risk aversion and return beliefs and, *as a result*, the individual's equity market participation.[7] Before examining the genetic components, we assess the relation between these characteristics and risk aversion, beliefs, and equity market participation. The results from these tests for equity market participation are fully consistent with previous work, e.g., equity market participation is positively related to trust and inversely related to growing up poor.

[7] The distinction between circumstances and the traditional economic factors (risk aversion and beliefs) can be blurred. For instance, Hong, Kubik, and Stein (2004) suggest that sociability may impact equity market participation as a result of either a non-traditional factor (i.e., an individual garners utility from talking about their stock market participation) or due to information costs. For instance, the authors suggest that an individual may revise their beliefs regarding the distribution of expected equity returns by learning via social relationships about the historically high stock return distribution. Thus, the "information cost" mechanism in this example is that socialization impacts beliefs regarding the distribution of expected equity returns.

Previous work that examines the channels through which these 11 variables impact equity market participation tends to focus on risk aversion, as (traditionally) data regarding heterogeneity in beliefs has been scarce. Our tests provide, often for the first time, evidence that most of these 11 explanatory variables are meaningfully associated with heterogeneity in *both* risk aversion and beliefs. For instance, individuals with greater trust tend to exhibit lower risk aversion, express a higher probability of a positive equity market return over the next year, and believe there is a lower probability of a greater than 20% decline in equity market returns over the next year.

We next investigate whether genetic scores predict heterogeneity in the 11 explanatory variables that, in turn, help explain heterogeneity in the traditional economic factors (risk aversion, beliefs, and equity market participation). Some of the links are obvious and direct—a genetic predisposition for Height should be associated with an individual's height, and taller individuals are more likely to participate in equity markets (Addoum, Korniotis, and Kumar (2017)). Even in less direct cases, however, intuition and previous work suggests a potential link between genetics and these mechanisms, e.g., Evans and Revelle (2008) suggest high neuroticism is associated with lower trust and Guiso, Sapienza, and Zingales (2008) find lower trust is associated with lower equity market participation.[8] Consistent with the explanation that molecular genetic endowments associated with cognition, personality, health, and body shape help determine these explanatory variables, each of the 11 variables is meaningfully related to at least three of the eight genetic scores (individually) we examine and the average variable is meaningfully related to 5.4 of the eight genetic scores. Moreover, the results are intuitive and largely consistent with phenotypic evidence—for example, an individual's self-rated health assessment is positively related to genetic endowments associated with Educational Attainment and General Cognition, but negatively related to genetic endowments associated with Neuroticism, Depressive Symptoms, Myocardial Infarction, Coronary Artery Disease, and BMI.

Given the 11 characteristics are related to both genetic endowments and the traditional economic factors (risk aversion, beliefs, and equity market participation), our final test estimates the extent to which genetic endowments contribute to the relation between the 11 explanatory characteristics and the traditional economic parameters. Specifically, we partition the 11 characteristics into portions predicted by the eight genetic endowments and the portions unexplained by the genetic scores. We then estimate the relation between the traditional economic factors (risk aversion, beliefs, and equity

[8] In a related manner, Cohn, Engelmann, Fehr, and Marechal (2015) consider an economic experiment with finance professionals and find evidence that fear is negatively related to risky asset investments. To the extent fearfulness is related to neuroticism and depression, behavioral links might be expected to share genetic underpinnings.

market participation) and the portion of the explanatory variable (e.g., trust) associated with the eight genetic endowments and the portion orthogonal to the genetic scores. The results reveal that much of the relation between economic outcomes and the explanatory variables is attributed to the portion of the explanatory variable associated with genetic endowments. For instance, the portion of trust that is associated with the eight genetic scores accounts for about 41% of the relation between stock market participation and the two trust components—trust predicted by the genetic scores and trust independent of the genetic scores. Analogously, the "genetic component" of trust accounts for about one-quarter of the relation between trust and risk aversion and about one-third of the relation between trust and an individual's perception of the likelihood of at least a 20% fall in equity prices.

Our results have a number of implications. As noted above, comparisons of monozygotic and dizygotic twins reveal evidence of a heritable component associated with risk aversion and stock market participation. Although twin studies reveal evidence that a phenotype is heritable, they cannot provide evidence regarding why a phenotype is heritable. That is, while twins studies help identify "how much," molecular genetics help identify "how"—in the present study, understanding the heterogeneous origins of the genetic portion of risk aversion, beliefs, and stock market participation. An analogy is moving from understanding that heterogeneity in life experience influences risk aversion to understanding why life experience influences risk aversion (e.g., living through a depression is associated with greater risk aversion). For instance, we find that the genetic score associated with Neuroticism is positively associated with both risk aversion and beliefs of a higher probability that markets will experience a loss in excess of 20% over the next year. Moreover, the genetic Neuroticism score is negatively associated with trust and optimism (and is also related to multiple other variables posited to explain stock market participation). That is, a genetic predisposition associated with Neuroticism can help explain why some individuals are more risk averse, why expectations of equity returns differ across individuals, why some individuals are less trusting, and individual's portfolio choices. Genetic components also have the potential to clarify the impact of environment on choice. For example, genetic factors may help to isolate the impact of how life experiences impact portfolio choices (cf., Malmendier and Nagel (2011), Bharath and Cho (2016)) directly and in conjunction with genetic predispositions.

Second, our work provides the first evidence that beliefs regarding the distribution of equity returns (1) have a genetic component and (2) are related to many of the variables known to explain stock market participation (e.g., trust). For example, some postulate that heritability of portfolio decisions may arise from heritability of risk aversion (e.g., Barnea, Cronqvist, and Siegel (2010)).

Although our evidence is consistent with this causation channel, our tests reveal that genetics are also associated with heterogeneity in return beliefs. Importantly, we consider not only individual expectations that the market will increase over the next year (which reveals little about one's beliefs regarding the shape of the distribution of expected returns), but also beliefs regarding the likelihood of a severe market loss (a loss greater than 20%) and a large market gain (a gain greater than 20%). Our results demonstrate that, on average, expectations regarding the distribution of expected equity returns are remarkably biased (consistent with the limited work that has examined this issue). Moreover, we show that these biases are related to many of the genetic scores we examine. Historically, the likelihood of a 20% or greater decline in U.S. equities over a 12 month period is about 6.6%; yet, on average, individuals estimate the probability at 30.3%. Further, biased individual expectations regarding the distribution of expected equity returns can help explain why both genetic endowments (e.g., the genetic score for Neuroticism) and the explanatory variables used in previous studies (e.g., trust) are related to market participation. That is, what may be an objectively irrational choice (e.g., zero investment in equities), is "rational" to the person whose, at least partially innate, neuroticism causes them to believe equity investments are both extremely risky and have a negative risk premium.

Third, our results add to the understanding of genetics and finance by examining a different sample and using a different approach than previous work in this area. As Beauchamp et al. (2011) stress, molecular genetics differs from previous work in that the effect of the genotype is examined directly rather than inferred by comparing the behavioral similarity among subjects with shared biological relationships. Moreover, we provide the first examination focusing on the relations between genetics, risk aversion, beliefs, and equity market participation for a broad representative sample of U.S. investors (the U.S. Health and Retirement Study dataset) versus previous tests that use the Swedish Twins Registry. Samples from different populations are important in genetics based research because, as detailed in the next section, a trait's "heritability" is not a biological constant—rather it varies from one population to another and over time. This may be especially important in contrast with a relatively homogenous culture such as Sweden. Campbell (2006) points out, for example, that all Swedish households were exposed to a national financial education campaign in the late 1990s as part of the Swedish pension system reformation (Guo (2005) and Barnea, Cronqvist, and Siegel (2010) provide additional discussion regarding the limitations of the Swedish data). Further, our more recent data allows us to obtain a more precise estimate of equity market participation since we now have information on retirement accounts, which was unavailable for much of the previous work.

Fourth, our work has policy implications. Genetic endowments do not imply intervention is useless or that environment has no effect. Rather, the relation between specific genetic endowments and individual characteristics can (1) guide policy in determining which mechanism may be productively influenced and (2) facilitate understanding of why individuals will likely respond differently to policy interventions. For instance, our tests demonstrate that equity market participation is positively related to genetic scores associated with Educational Attainment but negatively related to genetic scores associated with Neuroticism. Thus, the modifiable channel or intervention needed to encourage participation for two otherwise identical individuals who differ in genetic susceptibility to Neuroticism versus two individuals who differ in genetic endowments associated with Educational Attainment likely vary.

Fifth, our tests shed light on understanding why characteristics, such as trust, optimism, or growing up poor, are not only associated with stock market participation, but also vary across individuals. First, we demonstrate that most of the 11 explanatory variables are associated with *both* risk aversion and beliefs regarding the distribution of equity returns. Second, we find that all 11 of the characteristics (identified in previous work) have at least some genetic origins and therefore the relations between these characteristics and equity market participation reflect both environmental and innate biological components. Trust, for example, is likely influenced by both educational background and religion (see discussion in Guiso, Sapienza, and Zingales (2008)). However, our results suggest heterogeneity in trust is also related to heterogeneity in genes associated with Neuroticism, Educational Attainment, BMI, Depressive Symptoms, Coronary Disease, and Myocardial Infarction. Thus, for example, such genetic endowments can help in understanding how changes in environment impact trust and equity market participation (e.g., Giannetti and Wang (2016)).[9]

Finally, our work adds to our understanding of non-economic (behavioral finance) motivations. Cesarani, Johannesson, Magnusson, and Wallace (2012) point out that evidence of a genetic component associated with economic outcomes helps overcome the criticism that behavioral economics often lacks theory to explain heterogeneity across individuals. For instance, our results

[9] An emerging genetics literature explicitly recognizes interesting and complex passive, reactive, or active linkages between an individual's genetic endowment and the environment. In our context, an example of a passive gene-environment interaction might occur if a high intelligence parent was more likely to encourage a high intelligence child to invest in the equity market as an example of an interesting academic endeavor. An example of a reactive gene-environment interaction might occur if a healthy, non-neurotic, tall individual was more likely to receive positive feedback from society that in turn resulted in more optimistic distributional beliefs. Perhaps most important in our case with an aging sample, an example of an active gene-environment interaction could occur if positive distributional beliefs led to greater investment and then greater further investments later in life; whereas, less optimistic individuals might be less likely to begin any equity investment strategy.

tying genetic endowments associated with Neuroticism and Depressive Symptoms to risk aversion, return beliefs, and market participation, provide support for Caplin and Leahy's (2001) model that adds anxiety to the utility function. From a Darwinian perspective, our results provide empirical support for the notion that biological heterogeneity in decision making, although suboptimal for a given agent, may improve the likelihood of survival for a species (e.g., Brennan, Lo, and Zhang (2018)). Economic choices, health, and genetics may also be related. Our results support recently documented research identifying a relation between stock market losses and health outcomes (e.g., Cotti, Dunn, and Tefft (2015), McInerney, Mellor, and Nichols (2013)) and suggest a potential genetic component. For instance, the link between genetic scores (e.g., associated with Neuroticism and Depressive Symptoms) and traditional outcomes (such as risk aversion and beliefs regarding the distribution of expected equity returns) is consistent with a genetic explanation for the link between increases in California hospital admissions related to psychological conditions (such as anxiety and panic disorders) and reductions in California stock values (Engelberg and Parsons (2016)).

I. Background

A. Genetics and Finance – Twins Studies

In the last 50 years, at least 2,700 twin studies infer the role of genetics in more than 17,800 traits (see Polderman et al. (2015)) by comparing phenotype differences between monozygotic twins (who have, essentially, identical DNA) and dizygotic twins (who, the same as any other siblings, share approximately 50% of the DNA that varies across humans).[10] Although requiring "some strong assumptions" (Benjamin et al. (2012)), most twin studies use an established methodology to partition a phenotype's variance into three components—a genetic component, a shared environment component, and a non-shared environment component.[11] The genetic component estimate is denoted the phenotype's "heritability." Despite strong underlying assumptions, twin studies provide substantial information regarding the importance of genetics (see, e.g., the discussion in Winerman (2004)).[12]

[10] Most DNA (>99%) is identical across all humans. Thus, of the DNA that is not identical across humans, dizygotic twin siblings share approximately 50% (known as concordance).

[11] Specifically, twin studies require an "equal environment" assumption and no assortative mating (e.g., individuals with high cognitive skills are not more likely to mate with other high cognitive skill individuals). In addition, the methodology used in most twin studies is typically based on the assumption that gene effects are linear, there are no gene-gene interactions, and there are no gene-environment interactions (see Benjamin et al. (2012) for additional discussion). Joseph (2013) provides a well-known critique of twin study inferences in the social and behavioral sciences.

[12] Work in behavioral genetics employs a number of other methods to estimate the roles of heritability, shared environment, and non-shared environment including, for example, similarity in other related individuals (e.g., parent-child), reared apart twins, reared apart siblings, and adopted siblings. These alternative approaches largely reach the same conclusions (e.g., Turkheimer (2000)).

Although twin study results are often presented as "nature versus nurture," genetic endowments interact with environmental effects and therefore, heritability (usually reported as a point estimate) varies across populations and time due to these interactions. Guo (2005) points out, for example, that the twins study estimate of cognitive development heritability in a society that provides equal access to education will differ greatly from the estimate of cognitive development heritability in a society that has greatly imbalanced access to education (e.g., a population where only the wealthy have access to education).[13,14] The interaction of environment and genetic endowments is, of course, not limited to twin studies. Research in behavioral genetics also demonstrates that genes and environment usually interact (e.g., Dick (2011)).[15]

In recent years, a series of studies employing data from the Swedish Twins Registry estimate heritability for various financial phenotypes. Cesarini, Dawes, Johannesson, Lichtenstein, and Wallace (2009) estimate that approximately 20% of risk preferences are heritable. Cesarini, Johannesson, Lichtenstein, and Wallace (2009) estimate that 16-34% of overconfidence is heritable. Cesarini, Johannesson, Lichtenstein, Sandewall, and Wallace (2010) estimate that approximately 25% of portfolio risk choices are heritable. The authors also find a heritable component associated with choosing a default fund, choosing an ethical fund, and trend chasing. Cesarini, Johannesson, Magnusson, and Wallace (2012) find evidence that behavioral biases including the conjunction fallacy, default bias, and loss aversion are heritable.

Also using the Swedish Twins Registry data, Barnea, Cronqvist, and Siegel (2010) estimate that about one-third of the variation in equity market participation, fraction of wealth invested in equities, and portfolio volatility is heritable. Cronqvist and Siegel (2014) estimate that behavioral biases

[13] Vineis and Pearce (2011) also provide a related discussion regarding the problems with estimating heritability via the traditional twin study methodology. Specifically, phenylketonuria (PKU, a rare inherited blood disorder for which all U.S. newborns are tested), can be effectively avoided by: (1) not having the mutated gene, or (2) eating a low phenylalanine diet. If one studies a population where there is no variation in diet, 100% of the variation in outcome is due to genes. If one studies a population where all have the genetic mutation but there is variation in the diet, then 100% of the variation is due to diet. Thus, in contrast to classical twin studies that require the causes (genetic, shared environment, non-shared environment) to sum to 100%, PKU causes could sum to 200%, i.e., 100% of the cases could be avoided by not having the genetic mutation and 100% of the cases could be avoided by eating a low phenylalanine diet.

[14] Consider for example, that height is highly heritable (approximately 70%) yet in the last 150 years the average height in industrialized nations has increased about four inches due to environmental changes such as better nutrition.

[15] The partition of impacts into genetic and environmental components is always, at least in part, a function of the sample characteristics. For example, many educational attainment studies find little role for the environment relative to genetic variability in explaining educational outcomes. In stark contrast, studies of charter and Catholic schools often find evidence supporting the role of environment in impacting educational outcomes. Stoolmiller (1999) and Sacerdote (2011) stress that the resolution to the conundrum lies in the recognition that most genetic studies of educational attainment do not include sufficient variability in the environmental variables (the restricted range problem). In a related, but contrasting manner, limited heterogeneity in the genetic population may exacerbate the role of environmental forces.

including lack of diversification, excessive trading, and the disposition effect are highly heritable. Cronqvist and Siegel (2015) estimate that about one-third of the variation across savings rates is heritable in Sweden. Calvet and Sodini (2014) use twins in the Swedish registry to investigate the relation between wealth and risk aversion.

B. Molecular Genetics

The human genome—the genetic information needed to build and maintain a human—is contained in 23 pairs of chromosomes (46 total, 23 from each parent) within the nucleus of each human cell.[16] Each chromosome is a tightly packed DNA molecule of the familiar double helix shape. Genes are particular regions of DNA (humans have an estimated 20,000 genes) that "code" (i.e., contain the information needed to make) proteins.[17,18] In fact, most DNA (>98%) is so-called noncoding DNA. The "rungs" of the DNA "ladder" consist of pairs of four nitrogenous bases—adenine, thymine, cytosine, and guanine, typically referred to as A, T, C, and G, where A always pairs with T, and C always pairs with G (known as complementary Watson-Crick base pairings).[19] Focusing on one side of the DNA "ladder," a single nitrogenous base combined with the "rail" of the ladder (consisting of a sugar molecule and phosphate group) is denoted a nucleotide. Genome sequencing is the process of ordering these nucleotides in the genome. When referring to DNA sequencing, scientists denote the sequence along one side of the ladder. For example, a sequence may be ATTAGGC (where the first A will be paired with a T on the other side of the ladder, and so on). Although the human genome contains approximately 3 billion base pairs, only about 0.6% vary across individuals and account for genetic differences across individual humans.[20]

---

[16] Except red blood cells (that do not have a nucleus) and germline (sperm and egg) cells that contain 23 chromosomes.

[17] For simplicity we refer to genes as the coding regions of DNA (the "classic definition"). Advances in genomic research, however, reveal that genes include non-coding portions as well (e.g., introns). As a result, there is a lack of consensus on the exact definition of a gene. An alternative, for example, is (see Keller and Harel, 2007), "A gene is a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions and/or other functional sequence regions."

[18] Phenotypes are also influenced by environment within the body that can, in turn, be impacted by the environment outside the body. Specifically, epigenetics examines "gene expression," i.e., which genes are activated and to what extent. In a simple analogy, DNA is hardware and epigenetics is software. Epigenetics is impacted by many factors including nutrition, sleep, aging, and exercise. Evidence suggests epigenetic impacts can be heritable. In a landmark study, for example, Katti, Bygren, and Edvinsson (2002) found that when grandfathers had excess food supply during their "slow growth period" (typically ages 9-12) just prior to puberty, grandsons suffered from increased death from heart disease and diabetes.

[19] The two pairs A-T, C-G combined with the two "rails" of the DNA "ladder" to produce four possible combinations, i.e., A-T, T-A, C-G, or G-C.

[20] The human genome can have other differences including insertions (extra nucleotide base pairs), deletions (missing nucleotide base pairs), copy number variants (parts of the genome include repeating patterns that can vary in the number

The places where genetic sequencing differs across two individuals are known as single nucleotide variation (SNV). For instance, at a specific location within the genome (known as a locus or, plural loci), an A may be replaced with a C. When less than 1% of the population exhibits variation at a specific locus, the variation is known as a mutation. When at least 1% of the population exhibits the pattern (e.g., when a substantial portion of the population exhibits this variant), the location is denoted a Single Nucleotide Polymorphism (SNP, pronounced "snip" or, when referring to multiple SNPs, "snips"). SNPs occur, on average, about once every 300 nucleotides implying about 10 million SNPs in the human genome. The more common variation is denoted the major allele (e.g., adenine (A)) while the less common variation is the minor allele.[21] Because individuals have two of each chromosome (one from each parent), an individual can have 0, 1, or 2 of the minor allele.

Molecular genomic research primarily focuses on consortium-based meta-analysis, or Genome-Wide Association Study (GWAS).[22] This is a relatively new field—the first GWAS was published in 2005 (Klein et al. (2005)). The GWAS approach has proven highly successful: a recent review of the method in the *American Journal of Human Genetics*, concludes that "…the empirical results [associated with GWAS] have been robust and overwhelming…" (Visscher et al. (2017)). Broadly, a GWAS takes an atheoretical approach and examines the relation between the outcome (phenotype) and (roughly) the entire variation in the genome. This approach recognizes that most traits are associated with many loci that individually have small effects on the phenotype.[23] That is, in most cases, outcomes are related to hundreds or thousands of SNPs. For instance, although height is highly heritable, there is no "height

of repeats), and aneuploidy (extra or missing chromosomes). These variants are also associated with phenotypes, e.g., Down syndrome, a type of aneuploidy, results from three (rather than the usual two) copies of chromosome 21.

[21] Most SNPs are biallelic—meaning there are only two variations (e.g., either A or G in a sequence).

[22] Most GWAS are consortium based partnerships of multiple universities and research organizations that allow sharing of analysis for meta-analysis across samples (with pre-defined protocols) without sharing protected genotype or clinical information (see Bush and Moore (2012) for additional detail). Our discussion of meta-analysis consortium GWAS is necessarily brief and incomplete. For fuller descriptions see Visscher et al. (2017) and Evangelou and Ioannidis (2013).

[23] Early molecular genetics work employed the candidate gene approach in which a researcher proposes a theoretical argument as to why specific SNPs will be related to an outcome and then estimates a regression of the outcome (i.e., the individual's observable phenotype such as height) on the small set of SNPs. Unfortunately, the candidate gene approach is largely viewed as a failure as such studies suffer from an extremely high rate of false positives. Benjamin et al. (2012) provide an excellent summary of the pitfalls associated with the candidate gene approach. The authors reference a study by Obeidat et al. (2011), for example, that was only able to replicate results from one of 104 published candidate gene studies based on an independent sample. Quite possibly, our collective knowledge of the pathways between genetic endowments and diseases, or particular phenotypes, remains too rudimentary for this approach to be successful given the large number of candidate genes available. A related cause of this failure is that although some very rare diseases (such as sickle cell anemia) are associated with a single locus (i.e., a Mendelian trait), nearly all traits of interest to social scientists (e.g., neuroticism) and most medical conditions (e.g., heart disease) are complex and associated with many SNPs across many genes (as well as environmental factors and interactions between the two).

gene"—rather the GWAS used for height identifies 697 statistically significant SNPs across the genome.[24]

As a specific example of a GWAS, the Social Science Genetic Association Consortium (see Okbay et al. (2016)) investigated 9.3 million SNPs on a discovery sample of 293,723 individuals and a replication sample of 111,349 individuals to examine the relation between genes and educational attainment. Because the number of potential independent variables, 9.3 million, is much greater than the number of observations, researchers cannot estimate a multiple regression. As a result, a GWAS begins by regressing the phenotype on each individual SNP (and controls) and then weights the SNPs by effect size to form a single quantitative measure of the relation between genetics and a phenotype—this measure is known as a Polygenic Score (PGS, also known as polygenic risk score, genetic risk score, or genome-wide score) associated with the phenotype.

In an important study that touches on molecular genetics and finance, Barth, Papageorge, and Thom (2018) examine the relation between an Educational Attainment PGS and wealth inequality for participants in the 2006 and 2008 waves of the Health Retirement Study (HRS).[25] Although the focus of their study is understanding the relation between their Educational Attainment PGS and wealth inequality, the authors also find that the link between Educational Attainment and income inequality arises, in part, because their Educational Attainment PGS is inversely related to a measure of risk aversion, positively related to stock market participation, and positively associated with more accurate (i.e., smaller deviation from the historical average) beliefs regarding the likelihood of a positive market return. To the best of our knowledge, we are the first study to focus on the relation between molecular genetic endowments associated with cognition, personality, health, and body shape and traditional economic outcomes including equity market participation, risk aversion, and beliefs regarding the distribution of equity returns. We also differ from Barth, Papageorge, and Thom in that we examine: (1) a broad range of genetic endowments—cognition, personality, health, and body shape—that are plausibly associated with characteristics previously shown to impact equity market participation, and

[24] Because the number of examined SNPs is so large, to correct for the false positive problem, the statistical significance, or *p*-value, for a SNP is typically adjusted using a Bonferroni approach that requires rejection at the critical value associated with $p<5 \text{ X } 10^{-8}$. That is, the height GWAS identifies 697 SNPs with *p*-values less than 0.00000005. Quick interpretations of GWAS studies usually occur with a Manhattan plot where the vertical axis is given in $\log_{10}$ scale with a critical value of 7.3 ($=\log_{10}(5X10^{-8})$.

[25] Shin, Lillard, and Bhattacharya (2018) examine the relation between a PGS for Alzheimer's Disease and saving behavior. The authors conclude that the Alzheimer's Disease PGS does not impact savings behavior or asset allocation decisions after accounting for age. Brown and Sias (2019) examine the relation between technology adoption and the same PGSs we examine in this study.

(2) how those genetic endowments relate to other variables (e.g., trust) previously shown to help explain heterogeneity in equity market participation.

Although GWAS-based molecular genetics has become the focus of modern genetics research, interestingly, molecular genetics can only account for a fraction of heritability estimated via twin studies. Most work suggests that this "missing heritability" arises from failing to adequately capture the information contained in the genome (e.g., Girirajan (2017)), although some work suggest that at least part of the problem arises from mismeasuring heritability in twin studies (see, e.g., Zuk, Hechter, Sunyaev, and Lander (2012), Vineis and Pearce (2011)). Regardless, the missing heritability question is endemic to genetics research when comparing results from molecular genetics to results from twin studies.[26]

C. Risk Aversion, Expected Return Distribution Perceptions, and Equity Market Participation

Traditional economic theory holds that if investors are rational, utility is solely a function of wealth, and risk aversion determines the shape of the utility function, an individual's investment decisions are (Brennan and Lo (2011)), "…completely determined by utility functions, budget constraints, and the probability laws governing the environment." Moreover, given the historical distribution of equity returns, traditional economic theory finds that most individuals should hold nearly 100% of their savings in equity (e.g., Heaton and Lucas (1997)). Surprisingly, however, the reported equity participation rates for individuals in the U.S. over the last four decades have averaged about 31% when excluding IRAs, and 43% when including IRAs (Giannetti and Wang (2016)). Rates tend to be even lower for most European countries (e.g., Georgarakos and Pasini (2015)). As discussed in the introduction, the variation in equity market participation is usually attributed to heterogeneity in risk preferences, heterogeneity in the beliefs regarding the distribution of expected equity returns, or heterogeneity in circumstances such as trading frictions or uninsurable background risk.[27]

[26] A brief review of any recent edition of *Nature—Journal of Human Genetics* reveals that molecular genetics dominates recent genetics research.

[27] Because traditional theory suggests investors use the same model of expected returns and incorporate all available historical information when forming expectations (e.g., see discussions in Vissing-Jorgensen (2002) and Malmendier and Nagel (2011)), investors, theoretically, have identical expectations. Previous empirical work suggests, however, that a number of factors may contribution to heterogeneity in beliefs including for example, life experience (see Benartzi (2001), Kaustia and Knupfer (2008), Choi, Laibson, Madrian, and Metrick (2009), Malmendier, Tate, and Yan (2011), Malmendier and Nagel (2011, 2016), Cameron and Shah (2015), Bharath and Cho (2016), Giannetti and Wang (2016), Knupfer, Rantapuska, and Sarvimaki (2017), Bernile, Bhagwat, and Rau (2017), and Anagol, Balasubramaniam, and Ramadorai (2018)).

This equity non-participation (or under-participation) puzzle has led to a large literature investigating factors that explain heterogeneity in equity market participation. Consistent with individuals facing a fixed cost associated with investing in equities (e.g., Vissing-Jorgensen (2002)), stock market participation is positively related to wealth and income. Fixed costs, however, can explain only a small portion of the observed levels of non-participation (e.g., Grinblatt, Keloharju, and Linnainmaa (2011)). Advances in the past two decades demonstrate that equity market participation is associated with a number of additional factors including, for example, education, cognitive ability, trust, sociability, optimism, negative early life economic experiences, body shape, health, race, political preferences, political activism, knowing your neighbors, credit score, and ambiguity aversion.[28]

Moreover, evidence suggests investors engage in a number of presumably irrational behaviors (e.g., loss aversion). Several theoretical models suggest innate biological differences may help explain heterogeneity in investor behavior—including what appears to be irrational behaviors. Caplin and Leahy (2001) show that adding anxiety (research suggests anxiety is highly heritable, e.g., Torvik et al. (2016)) to the utility function can help explain time inconsistency in preferences, the equity premium puzzle, and the risk-free rate puzzle. In a series of papers, Brennan and Lo (2011, 2012), Zhang, Brennan, and Lo (2014a, 2014b), and Brennan, Lo, and Zhang (2018) suggest that biologically generated variation in economic behaviors could be suboptimal for the individual (i.e., the individual acts "irrationally"), but optimal for the species, as diversity in behaviors can help ensure species success in the case of rare environmental events.

II. Data

A. Health Retirement Study

Our data comes from the Health and Retirement Study (HRS) survey data of a panel of more than 20,000 Americans age 50 and older.[29] This age cohort is important in financial markets as we estimate that individuals over 55 account for at least 74% of the value of stocks and 64% of the financial assets

[28] As noted above, the equity market participation literature is very large (Google Scholar reports 8,250 results for "stock market participation" in January 2019). The first 11 variables in this list are used in this study and discussed in the next section. We lack sufficient data for most of the other variables shown to help explain stock market participation. For evidence regarding the other variables mentioned—race, political preferences, political activism, knowing your neighbors, credit score, and ambiguity aversion—see, respectively, Vissing-Jorgensen (2002), Kaustia and Torstila (2011), Bonaparte and Kumar (2013), Brown, Ivkovic, Smith, and Weisbenner (2008), Bricker and Li (2017), and Dimmock, Kouwenberg, Mitchell, and Peijnenburg (2016).

[29] The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan.

held by individuals in 2016.[30] The HRS surveys were first administered in 1992 and then every two years (HRS interview "waves") through 2016. Moreover, the sample increases over time to add new respondents (HRS "cohorts"). For example, the most recent cohort (for which we have data) includes individuals born between 1954 and 1959 (the "Mid Baby Boomers") while the previous cohort (the "Early Baby Boomers") includes individuals born between 1948 and 1953. We limit our sample to respondents between ages 50 and 80 (average age of 67). In addition, as detailed below, questions identifying an individual's beliefs regarding the distribution of expected stock returns did not begin until 2010. Thus, our sample includes four HRS waves—2010, 2012, 2014, and 2016. Beginning in 2006, HRS selected one-half of the interviewee households (in that wave) for an "Enhanced Face-to-Face Interview" that included saliva collection (the raw source for the genetic data) as well as a "Leave-Behind Questionnaire" that included "Psychosocial and Lifestyle" questions. The 50% random enhanced face-to-face sample is rotated in every wave, i.e., Leave-Behind Questionnaire respondents in 2006 are also given enhanced face-to-face interviews in 2010 while those not selected in 2006 are selected for 2008 (and again for 2012, and the sample includes any new cohorts added).

For households consisting of partners (e.g., husband and wife), both partners are questioned. Only one person (the household's "financial respondent"), however, answers the household financial data questions. Although some variables are measured at the household level (e.g., wealth, stock market participation), genetics, risk aversion, beliefs regarding the distribution of expected equity returns, and the most of the 11 explanatory variables used in previous work (e.g., trust) are measured at the individual level. Therefore, we limit our analysis to financial respondents. Nonetheless, in the Internet Appendix, we repeat our tests incorporating genetic endowments for both spouses when examining (household) stock market participation and find similar results.

One concern is whether surveys adequately capture individual characteristics, beliefs, and behaviors. For example, in assessing individual beliefs regarding the likelihood of a positive or negative stock market return, individuals tend to report round numbers (e.g., 30% rather than 29% or 31% chance the market will fall more than 20%) when estimating probabilities. Similarly, measures of individual characteristics, such as trust, are generated via responses to a series of questions regarding trust (as one cannot view "trust"). Nonetheless, because most of the variables used to help explain heterogeneity in equity market participation are measured with error (e.g., trust, sociability), work in

[30] Using the Survey of Consumer Finances (https://www.federalreserve.gov/econres/scfindex.htm), these figures are based on estimates (Tables 6-89 and 1-01-16) for individuals age 55 and older (i.e., the SCF age breakpoint).

this area is necessarily based on less than perfect measures.[31] Moreover, the measurement error in our data should only serve to weaken the power of the tests.

B. Molecular Genetic Data

To measure the genetic predisposition for each individual to develop a particular trait, we employ the polygenic scores (PGSs) computed by HRS:

$$PGS_i = \sum_{j=1}^{J} W_j G_{i,j} \quad (1)$$

where PGS_i is the polygenic score (for a given phenotype) for individual i, W_j is the consortium meta-analysis GWAS weight (based on the odds ratio or beta estimates from the GWAS) for SNP j, and $G_{i,j}$ is HRS individual i's genotype (i.e., number of reference alleles (0, 1, or 2)) for SNP j.[32] In most cases, the GWAS weights are computed without HRS data, i.e., the HRS PGSs are out of sample estimates.[33,34] HRS scales the PGSs to zero mean and unit variance.

Because most of the GWAS PGS weights are based on European ancestry samples, and PGSs may not be directly applicable across ancestry groups (e.g., Martin et al. (2017)), we limit our sample to individuals of European ancestry. In addition, even within a given general ancestry, "population stratification" can contaminate results. Population stratification occurs when individuals in the same sample differ in SNP frequency, e.g., because SNPs are hereditary, Southern Europeans may exhibit a SNP more often than Northern Europeans (the root cause is non-random mating usually driven by physical separation). This can lead to confounding false positives when environmental factors impact

[31] Some work is based on samples that include relatively complete financial data. For instance, Grinblatt, Keloharju, and Linnainmaa (2011) examine the relation between cognitive ability and equity market participation for Finnish males where the authors have tax records to infer ownership and mandatory IQ tests (an estimate of cognitive ability) to ensure the sample does not suffer from a self-selection bias.

[32] Beauchamp et al. (2011) point out that both theory and evidence from animal breeding and behavioral genetics supports the use of the simple linear additive model that has become standard in molecular genetics research.

[33] HRS participates in several of the GWAS consortiums. In almost all of these cases, however, HRS data are excluded prior to generating the weights used to compute the HRS PGSs. Per correspondence with HRS, the PGS associated with BMI is the only PGS that includes the HRS sample in computing the GWAS weights used by HRS. The HRS sample, however, contributes less than 3% of the BMI observations (see Ware, Schmitz, Gard, and Faul (2018)).

[34] We use the PGSs computed by HRS. All PGSs are a function of the sample and weights computed by the consortium (e.g., better PGSs for a given phenotype will be found over time as the consortium sample size increases). Moreover, computation of the PGS requires the researcher (in this case, HRS) to make a number of decisions regarding the method, e.g., whether to include only highly statistically significant SNPs, or all SNPs. HRS, for example, includes all available SNPs in their estimate as recent work (e.g., Simonson, Willis, Keller and McQueen (2011), Abraham, Kowalczyk, Zobel, and Inouye (2013), Abraham and Inouye (2014), Goldstein, Yang, Salfati, and Assimes (2015), Abraham et al. (2016), Ware et al. (2017)) suggests that doing so produces better (out of sample) predictions (helping to mitigate the missing heritability issue). Because PGSs depend on both the consortium sample and the construction method, our HRS computed Educational Attainment PGS is different than the Educational Attainment PGS used by Barth, Papajohn, and Thom (2018).

phenotypes. The most popular way to solve this issue is to include SNP principal components in the analysis that, theoretically, capture within sample variation in SNPs. HRS provides the first 10 principal components for the HRS European ancestry sample. Following the guidance in Ware, Schmitz, Gard, and Faul (2018), and as detailed below, we include all 10 HRS SNP principal components in our tests.

We use the April 2018 version (HRS v2) of the HRS PGS data that incorporates DNA from HRS participants in the extended face-to-face samples from 2006-2012. The data include PGSs for 29 phenotypes. We focus on eight PGSs based on previous evidence of relations between individual characteristics and equity market participation.[35] Because risk aversion and beliefs are complex characteristics, they are likely related to multiple genetic endowments. For instance, although there is no risk aversion PGS (or GWAS), we expect risk aversion will be positively related to genetic endowments associated with both Neuroticism and Depressive Symptoms.[36] As a result, we select two PGSs from each of the four broad genetic classifications we examine (cognition, personality, health, and body shape). First, given evidence that cognition is positively related to equity market participation (e.g., Kezdi and Willis (2003), Benjamin, Brown, Shapiro (2013), Christelis, Jappelli, and Padula (2010), Grinblatt, Keloharju, and Linnainmaa (2011), Cole, Paulson and Shasty (2014)), we include the Educational Attainment PGS and the General Cognition PGS in our analysis. Second, given personality traits such as trust, sociability, and optimism are associated with equity market participation (e.g., Hong, Kubik, and Stein (2004), Puri and Robinson (2007), Guiso, Sapienza, and Zingales (2008), Heimer (2014), Balloch, Nicolae, and Philip (2015), Giannetti and Wang (2016)), we include the Neuroticism PGS and the Depressive Symptoms PGS. Third, given health is positively associated with equity market participation (e.g., Rosen and Wu (2004), Yogo (2016)) we include the Myocardial Infarction PGS and the Coronary Artery Disease PGS (heart disease is the leading cause of death in the U.S.). Fourth, given evidence body shape is associated with participation (BMI is negatively associated with market participation and height is positively associated with participation, Addoum, Korniotis, and Kumar (2017)), we include the BMI PGS and the Height PGS. Appendix A provides details regarding the construction of the HRS PGSs.

[35] For completeness, the Internet Appendix provides our primary tests for an additional 21 PGSs.
[36] See Kamstra, Kramer, and Levi (2012) for discussion of the evidence linking risk aversion to depression and Nicholson, Soane, Fenton-O'Creevy, and Willman (2005) for evidence linking risk aversion to neuroticism.

C. Outcomes – Risk Aversion, Expected Return Distribution Perceptions, and Stock Market Participation

We focus on risk aversion, return beliefs, and equity market participation as the traditional economic outcome variables. We measure risk aversion through the following question asked in the 2014 and 2016 HRS waves: "Are you generally a person who tries to avoid taking risks or one who is fully prepared to take risks? Please rate yourself from 0 to 10, where 0 means 'not at all willing to take risks' and 10 means 'very willing to take risks.'" To proxy for an individual's risk aversion across the 2010-2016 sample period, we average each respondent's score to this question over years 2014 and 2016 and subtract the average from 10 (so that higher values indicate greater risk aversion). As detailed below, we use a standardized (i.e., rescaled to zero mean, unit variance) version of this measure in our empirical tests for ease of interpretation.[37]

We use three questions to gather information regarding respondent beliefs about the distribution of expected returns. Starting in 2002, the HRS asked respondents, "We are interested in how well you think the economy will do in the future. 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?"[38] Beginning in 2010, HRS added two questions regarding the distribution of expected stock returns. Specifically, the first question asks, "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 have gained in value by more than 20 percent compared to what they are worth today?" The second question asks the probability for the other side of the distribution, i.e., the chances stocks "…have fallen in value by more than 20 percent…?"[39]

---

[37] As detailed in the Internet Appendix, we also consider a measure of relative risk aversion (inferred from "income-gamble" questions asked in earlier waves) and find similar results. We focus on self-rated risk aversion because (1) research suggests such questions better capture risk aversion (Guillemette, Finke, and Gilliam (2012)) relative to income gamble questions, and (2) the income gamble questions are asked between 1998 and 2006 and therefore exclude 40% of our sample.

[38] HRS has a section devoted to expectations. To ensure respondents understand the meaning, the section is introduced by "Next we would like to ask your opinion about how likely you think various events might be. When I ask a question, I'd like for you to give me a number from 0 to 100, where '0' means that you think there is absolutely no chance, and '100' means that you think the event is absolutely sure to happen. For example, no one can ever be sure about tomorrow's weather, but if you think that rain is very unlikely tomorrow, you might say that there is a 10 percent chance of rain. If you think there is a very good chance that it will rain tomorrow, you might say that there is an 80 percent chance of rain."

[39] One concern regarding probability questions is that individuals sometimes answer 50% to convey a lack of confidence in their response (see Fischhoff and Bruine de Bruin (1999), Lillard and Willis (2001)). During our sample period, respondents who answered 50% to the question regarding the likelihood the market will increase in the next year, were asked a follow up question of whether their 50% answer meant that they believed "it is about equally likely that these mutual fund shares will increase in worth as it is that they will decrease in worth by this time next year, or are you just unsure about the chances." For those that answer they are "just unsure," HRS does not ask the latter two questions regarding the likelihood of a greater than 20% decline and a greater than 20% increase. Because our sample is limited to individuals with equity market participation, risk aversion, and beliefs data, we exclude these individuals from our main

We generate two measures of stock market participation—an indicator for holding equity (either directly or in retirement/IRA/Keogh accounts) and the fraction of financial wealth (both directly held and in retirement/IRA/Keogh accounts) in equities.[40] Note that given the structure of the HRS interviews, in some cases we can observe if the respondent participates in equity markets, but cannot estimate the fraction of wealth invested in equity.[41] Appendix A provides details of the construction of these measures.

D. Other Variables used to Explain Stock Market Participation

We consider 11 variables previous work suggests can help explain equity market participation (via their impact on risk aversion, beliefs, or circumstances): wealth, income, education, cognitive ability, trust, sociability, optimism, negative early life economic experiences, height, BMI, and health. Although a number of previous studies use HRS data, the survey has evolved over time, and because we focus on recent data our measures are often more direct than earlier measures, e.g., more recent surveys include psychosocial questions directly focused on measuring trust, sociability, and optimism. Appendix A provides details regarding the construction of each of these 11 variables as well as a brief description of the proposed mechanism (from previous studies) linking the variable to risk aversion, beliefs, and equity market participation (and at least one associated reference).[42] We also include control variables in our analyses. Specifically, we include indicators for HRS waves, respondent's age, gender, retired, and married. Appendix A also provides details regarding construction of the control variables.

E. Descriptive Statistics

We require data for the respondent's genetics, risk aversion, beliefs regarding the distribution of equity market returns, stock market participation, household wealth, household income, age, gender,

analysis. As shown in the Internet Appendix, however, our results are essentially identical when including these individuals in the sample.

[40] Because much of the work in equity market participation focuses on earlier surveys that lack sufficient data to infer equity holdings in retirement funds (pension/IRA/Keogh), as a robustness test, we repeat our analysis focusing only on direct holdings (i.e., excluding retirement funds). Results, reported in the Internet Appendix, are essentially unchanged.

[41] For example, a respondent may report they have direct holdings in equity, but when queried about the value, they may respond they do not know, or refuse to answer.

[42] Although Appendix A provides discussion of these links for each variable, several themes are clear. For instance, Malmendier and Nagel (2011) propose that one's early life experience may influence both risk aversion and beliefs, thereby influencing equity market participation. This mechanism can be applied to all the variables consider, e.g., optimism may be positively associated with both risk tolerance and beliefs regarding the distribution of equity returns, thereby impacting equity market participation. We examine the relation between the traditional explanatory variables (e.g., early life poverty), risk aversion, beliefs, and equity market participation in Section IV.

marital status, and retirement status to be included in the sample. Our final sample consists of 5,560 individuals and 12,633 individual-year observations over four HRS waves (2010, 2012, 2014, and 2016). Panels A, B, C, and D in Table I report descriptive statistics for the genetic data (PGSs), economic outcome variables, the 11 explanatory variables used in previous studies, and the control variables, respectively. Panels A, B, and C in Table II report, respectively, correlations for the genetic data, economic outcome variables, and the 11 explanatory variables used in previous studies. Note that although some of the variables, such as PGSs, are observed at the individual respondent level and others, such as trust, are based on a respondent's estimated average trust over a period of time (see Appendix A), the values in Tables I and II are based on the pooled sample used in our analysis, i.e., there are 5,560 unique Neuroticism PGS observations and the average respondent's equity market participation and beliefs regarding the distribution of expected equity returns is measured 2.3 times implying 12,633 total observations.

[Insert Tables I and II about here]

Panel A in Tables I and II report descriptive statistics and correlations for the genetic data. By construction, the HRS PGS scores are standardized.[43] As expected, the "pairs" of PGS measures are strongly related, e.g., the correlation between the General Cognition PGS and the Educational Attainment PGS is 0.27.

Panel B in Table I reveals that 62% of respondents in our sample hold some equities either via direct holdings or in their retirement accounts.[44] The average household in our sample has 38% of their financial assets in equities (last row of Panel B). The first row of Panel B reveals substantial variation in self-rated risk aversion with a standard deviation of 2.04 on a scale of 0 to 10. As noted above, we use a standardized version (second row) of this variable throughout the empirical analysis for ease of interpretation.

The third through fifth rows in Panel B report respondent beliefs regarding the likelihood that the market will increase over the next year, the chance it will increase by more than 20% over the next year, and the chance it will fall by more than 20% over the next year. Although a broad literature

[43] The PGSs provided by HRS are standardized. Because, however, not all individuals included in the HRS genetic dataset have sufficient information to be included in our sample we re-standardize (rescale to zero mean, unit variance) the HRS genetic data for our sample to ensure the mean and variance are exactly zero and one, respectively.

[44] We find that 31% of respondents hold equity directly (i.e., excluding retirement funds) consistent with estimates reported in earlier studies using the survey questions on direct holdings (e.g., Addoum, Korniotis, and Kumar (2017) report 29% of HRS households hold equity based on the 1992-2008 HRS waves). Consistent with recent evidence (e.g., Giannetti and Wang (2016)), the results demonstrate that equity market participation estimates are substantially greater (at least in recent years) when including the additional information (often not available in earlier waves) regarding equities held in retirement accounts.

examines equity participation and a number of studies also examine risk aversion, comparatively little work focuses on understanding heterogeneity in individual beliefs regarding the distribution of expected equity returns. This results because, prior to our sample period, few surveys questioned individuals regarding their beliefs (especially perceptions related to the likelihood of a gain or loss greater than 20%). Consistent with previous work, however, the values reported by respondents differ greatly from the historical distribution of equity returns.[45] Specifically, between 1927 and 2016, rolling 12-month U.S. equity returns averaged 11.86% with a standard deviation of 21.29%. Nearly identical to the expected values under a normal distribution, 73.9% of the historical annual returns were positive, 33.8% were greater than 20%, and 6.6% of the observations were less than -20%. The typical individual, however, is much too bearish relative to historical norms—the average individual *greatly* underestimates the likelihood of the market increasing (mean estimate of 47% versus 74% historically) and greatly overestimates the likelihood of a 20% or greater decline in stock prices (mean estimate of 30% versus less than 7% historically). Individuals, on average, are closer to the historical value for the likelihood of a 20% or greater market gain (mean estimate of 28% versus 34% historically).

Panel B of Table II reports the correlation between the economic outcome variables. Consistent with theory, equity market participation (last two rows) is inversely related to risk aversion and positively related to an individual's beliefs regarding the likelihood that equity markets will rise in the next year. Market participation is inversely related risk aversion and an individual's beliefs regarding the likelihood of a greater than 20% decline in equity values but positively related to beliefs that market will increase in the next year. Consistent with recent work (Lee, Rosenthal, Veld, and Veld-Merkoulova (2015)), risk aversion is inversely related to individual's beliefs regarding the expected likelihood of a positive market return over the next year. Although the likelihood of a 20% gain is positively related with the likelihood of the market increasing, it is largely independent of risk aversion and only weakly related to market participation (in these pairwise correlations).[46]

Panel C of Tables I and II reports descriptive statistics and correlations, respectively, for the 11 explanatory variables used in previous studies. Table II reveals that many of the factors known to

[45] Several studies find that individuals tend to underestimate the likelihood of the market increasing in the next year (e.g., Dominitz and Manski (2007, 2011), Hurd and Rohwedder (2012), Hurd, van Rooij, and Winter (2011)). Goetzmann, Kim, and Shiller (2017) find that both individual and institutional investors greatly overestimate the likelihood of an extreme one-day stock market crash (that the authors define as greater than 12.82%, i.e., the value of the largest negative one day return associated with the 1929 stock market crash).

[46] In later analyses, with the inclusion of control variables and 10 principal components, the marginal effect associated with the perceived likelihood of a greater than 20% increase in stock prices suggests it captures the respondent's view of the riskiness of stock investments.

explain equity market participation are strongly related. For example, consistent with Guiso, Sapienza, and Zingales (2008), trust is positively associated with sociability and consistent with Hong, Kubrik, and Stein (2004), sociability is positively related to education. As shown in Panel D of Table I, our respondents are approximately equally likely to be male or female, 61% are married, and the average age is 66.

III. Genetics, Risk Aversion, Beliefs, and Equity Market Participation

We begin by examining if the eight molecular genetic endowments help explain variation in the economic outcome variables. Our hypotheses are based on extant literature (see Panel C of Appendix A for details including a brief description of the proposed channels and associated references) that relates heterogeneity in individual characteristics to heterogeneity in risk aversion, beliefs, and equity market participation. First, given evidence that education and cognitive ability are positively associated with equity market participation, we hypothesize that PGSs associated with Educational Attainment and General Cognition will be associated with lower risk aversion, a higher perception of the probability of a positive market return, a lower perception of the likelihood of a 20% or greater decline in prices, and greater stock market participation. Although individuals, on average, underestimate the likelihood that the market will increase by more than 20%, they also greatly overestimate the likelihood of a negative market return and the likelihood of a greater than 20% decline in stock prices. Thus, we expect higher Educational Attainment and General Cognition PGSs will be associated with a more realistic view of the standard deviation of expected returns and, therefore, a lower perceived likelihood of a greater than 20% increase in equity prices.

Second, given evidence that trust, sociability, and optimism are positively associated with equity market participation, we hypothesize that PGSs associated with Neuroticism and Depressive Symptoms will be associated with higher risk aversion, lower perceived probability of a positive market return, higher perceived probabilities of a greater than 20% decline or rise in equity prices, and decreased market participation. Similarly, given evidence health is positively associated with equity market participation, we expect the PGSs associated with poor health—Myocardial Infarction and Coronary Artery Disease—will be associated with higher risk aversion, lower perceived probability of a positive market return, higher perceived probabilities of a greater than 20% decline or rise in equity prices, and decreased market participation. Last, given evidence that lower BMI and greater height are positively related to equity market participation, we predict that a smaller BMI PGS and larger Height

PGS will be associated with lower risk aversion, higher expectations of a positive market return, lower expectations of a greater than 20% decline or rise in equity prices, and greater market participation.

Panel A in Table III reports the results of panel regressions of the economic outcome variables on the control variables (indicators for HRS waves, age, gender, retired, and married; unreported to conserve space), the 10 HRS genetic principal components (unreported to conserve space), and each of the eight PGSs individually (standard errors are clustered at the respondent level). That is, Panel A reports the results from 48 different regressions (six outcomes times eight PGSs). Given the PGSs are standardized, the coefficients in Table III reflect the relation between a one standard deviation change in the PGS and the outcome variable. For instance, the top cell in the fifth column reveals that a one standard deviation increase in the Educational Attainment PGS is associated with a 6.5% higher likelihood (statistically significant at the 1% level) that an individual holds any equity.[47]

[Insert Table III about here]

The results in Panel A of Table III provide very strong support for our hypotheses and reveal that the eight genetic endowments we investigate are associated with risk aversion, investor beliefs regarding the distribution of expected equity returns, and equity market participation. Specifically, 44 of the 48 reported coefficients have the expected sign and, of those 44 coefficients, 31 differ materially from zero at the 10% level and 27 differ from zero at the 5% level (two-tail tests). None of the four coefficients with an unexpected sign differ materially from zero at the 10% level.

Although the PGSs are correlated (e.g., Panel A of Table II shows the correlation between the Neuroticism PGS and the Depressive Symptoms PGS is 0.55), we next include all the PGS (and the control variables and 10 genetic principal components) in a single regression for each of the economic outcome variables. Panel B in Table III reports the results. Panel A of Table III shows that when considering risk aversion and including each of the genetic factors by themselves as an independent variable, almost all were statistically significant. Panel B shows that when all of the PGS are included in the same regressions, some of the PGS have stronger effects than others. For example, risk aversion is most strongly related to the Neuroticism PGS—a one standard deviation greater Neuroticism PGS is associated with a 7.2% standard deviation higher risk aversion (recall the risk aversion measure is scaled to unit standard deviation). Nonetheless, even when accounting for an individual's genetic

[47] We focus on linear probability models for ease of interpretation, to facilitate comparison with much of the literature in this area (e.g., Barth, Papageorge, and Thom (2018), Giannetti and Wang (2016), Puri and Robinson (2007), Hong, Kubik, and Stein (2004)), and for our later focus on comparing the relative contributions from genetic and non-genetic components. In the Internet Appendix, we repeat our tests with limited dependent variable models (binary logit and fractional logistic) and find essentially identical results.

predisposition for Neuroticism, risk aversion is positively related to the Coronary Artery Disease PGS (at the 5% level).

The multiple regression (Panel B) analysis also reveals that individual beliefs regarding the distribution of equity returns are related to a number of genetic factors. Beliefs regarding the likelihood of a market increase are positively related to the Educational Attainment PGS (at the 1% level) and the Coronary Artery Disease PGS (at the 10% level). The latter is inconsistent with our priors. Consistent with our hypothesis, however, the Myocardial Infarction PGS is negatively related (statistically significant the 1% level) to beliefs regarding the likelihood of the market rising in the next year. Specifically, a one standard deviation higher Myocardial Infarction PGS is associated with approximately a 1.06% lower value for beliefs regarding the likelihood of the market increasing. Given the standard deviation of the dependent variable is 26.53% (see Table I), this represents a 4.0% standard deviation shift (i.e., 1.063/26.525=0.040).

The multiple regression analysis (Panel B) also reveals that views regarding the likelihood of an extreme market move are most strongly related to the Neuroticism PGS. Specifically, individuals with a one standard deviation larger Neuroticism PGS predict a 1.02% greater likelihood of market decline in excess of 20% (statistically significant at the 1% level) and a 0.73% greater likelihood of a market increase in excess of 20% (statistically significant at the 5% level) over the next year. Relative to the dependent variable standard deviations reported in Table I, the coefficients suggest that a one standard deviation larger Neuroticism PGS is associated with a 4.40% standard deviation higher likelihood of a greater than 20% decline and a 3.23% standard deviation higher likelihood of a greater than 20% increase. Beliefs regarding the likelihood of an extreme market movement are also related to the Educational Attainment PGS and the General Cognition PGS. Specifically, individuals with a higher General Cognition PGS tend to believe the distribution of equity returns is less likely to experience a greater than 20% decline (statistically significant at the 5% level) or a greater than 20% increase (marginally significant at the 10% level) in the next year. Individuals with a higher Educational Attainment PGS tend to believe the distribution of equity returns is less likely to experience a greater than 20% decline (marginally significant at the 10% level).

Within the multiple regression framework, equity market participation (last two columns of Panel B) is most strongly related to the Educational Attainment PGS. Even when controlling for the genetic endowment related to Educational Attainment, however, the Neuroticism PGS, the Depressive Symptoms PGS, the Myocardial Infarction PGS, and the BMI PGS all impact equity market participation. That is, each of the four genetic component groups we examine—cognition, personality,

health, and body shape—appears to impact equity market participation. Specifically, higher Educational Attainment PGS, lower Neuroticism PGS, lower Depressive Symptoms PGS, lower Myocardial Infarction PGS, and lower BMI PGS are all associated with greater equity market participation.

A. Market Participation and the Genetic Components of Risk Aversion and Beliefs

As discussed above, traditional economic theory holds that the proportion of equity in an investor's portfolio is a function of the investor's beliefs and risk aversion. Thus, we next estimate the role of the eight genetic endowments in explaining the relation between equity market participation, risk aversion, and beliefs regarding the distribution of expected equity returns. Our approach is to first remove variation in the outcome variables related to the control variables and genetic principal components and then partition the remaining variation into genetic (predicted) and non-genetic (residual) components by regressing the remaining variation in the outcome variables on the PGSs.[48] Because the predicted and residual components are, by construction, independent, the R^2 from the second regression can be directly decomposed into the portion attributed to genetic and non-genetic factors. Specifically, we begin by regressing the six outcome measures (the risk aversion metric, the three belief metrics, and the two equity market participation measures) on the control variables (indicators for HRS waves, age, gender, retired, and married) and the first 10 principal components of the genetics data. The residuals from these regressions—denoted orthogonalized outcomes—reflect variation in each outcome that cannot be explained by the control variables or the 10 genetic principal components.

To examine the relation between stock market participation, risk aversion, and beliefs, we regress orthogonalized stock market participation on orthogonalized risk aversion and orthogonalized beliefs. For ease of interpretation, we standardize (rescale to unit variance and zero mean) the independent variables. Panel A in Table IV reports the estimated coefficients from regressions of the two measures of orthogonalized equity market participation on orthogonalized risk aversion and beliefs individually (columns one to four and six to nine) and simultaneously (columns five and ten). Consistent with traditional economic theory, equity market participation is negatively related to risk aversion, positively

[48] We take this approach so that we can decompose the R^2 into genetic and non-genetic components. Recognize, however, that removing variation due to the control variables and the genetic principal components before regressing the outcomes on the PGSs biases our results away from the PGSs explaining the variation in the outcome variables (i.e., our estimates are conservative). In the Internet Appendix we report regressions including the controls, principal components, and PGSs simultaneously and find nearly identical results (although such an approach does not allow a direct decomposition of the R^2 into genetic and non-genetic components).

related to beliefs regarding the likelihood the market will increase in the next year, and inversely related to beliefs regarding the likelihood of a greater than 20% decline in equity markets over the next year regardless of whether the regressors are included individually or simultaneously. When included as the only regressor, the orthogonalized likelihood of a 20% gain is positively related to the equity market participation (columns three and eight). However, consistent with the hypothesis that investors who view the market as less risky are more likely to participate, when controlling for risk aversion and other beliefs, investors who believe there is a higher likelihood of a greater than 20% equity return in the next year are less likely to invest in equity markets (columns five and ten).

[Insert Table IV about here]

We next regress orthogonalized risk aversion and orthogonalized beliefs on the eight PGSs (regression estimates are reported in the Internet Appendix). We denote the portion of orthogonalized risk aversion or beliefs explained by the eight PGSs (i.e., the fitted values) as the genetic component of risk aversion or beliefs and the portion unexplained (i.e., the residuals) as the non-genetic portion of risk aversion or beliefs. We then regress orthogonalized equity market participation on the standardized (for ease of interpretation) genetic and non-genetic components of risk aversion and beliefs. For example, the panel regression of orthogonalized stock market participation (denoted EQ) on the orthogonalized portion of risk aversion associated with the eight PGSs (denoted RA_{PGS}) and the orthogonalized portion of risk aversion independent of the eight PGSs (denoted $RA_{Non\text{-}PGS}$) is given by:[49]

$$EQ = \beta_1 RA_{PGS} + \beta_2 RA_{Non-PGS} + \varepsilon \tag{2}$$

Because the two components are mechanically uncorrelated, the R^2 can be directly decomposed into the genetic and non-genetic components:[50]

$$R^2 = \frac{\sigma^2(\widehat{EQ})}{\sigma^2(EQ)} = \frac{\sigma^2(\hat{\beta}_1 RA_{PGS})}{\sigma^2(EQ)} + \frac{\sigma^2(\hat{\beta}_2 RA_{Non-PGS})}{\sigma^2(EQ)}. \tag{3}$$

Panel B of Table IV reports coefficients from regressions of the two measures of orthogonalized equity market participation on the genetic and non-genetic components of orthogonalized risk aversion and beliefs. The bottom row reports the fraction of the R^2 accounted for by the portion of

[49] We exclude the individual i and year t subscripts for notational brevity.

[50] For simplicity, we write Equations (2) and (3) when using only the genetic and non-genetic components of risk aversion. However, Equation (3) holds even when Equation (2) contains eight regressors, i.e., the four measures (risk aversion and the three belief metrics) predicted by the PGSs and the four residuals. That is, given the same set of predictors (i.e., the eight PGSs), each genetic component is orthogonal to its own residual and each of the other residuals. For instance, the genetic (i.e., predicted) portion of risk aversion is orthogonal to the non-genetic (i.e., residual) portion of the chance the market will fall by 20%.

risk aversion or beliefs predicted by the eight PGSs, i.e., the first term on the right-hand side of Equation (3) divided by the R^2.

The results in Panel B suggest that a substantial portion of the relation between the equity market participation and risk aversion or beliefs arises from the portion of risk aversion or beliefs associated with the eight genetic endowments. For example, focusing first on the regressions that include only the genetic and non-genetic components of a single variable (i.e., columns one to four, and six to nine), all coefficients (both the genetic portion and the non-genetic portion) are statistically significant at the 1% level and economically meaningful. For instance, the first column reveals that a one standard deviation higher "genetic" risk aversion is associated with a 4.9% lower chance of holding equity while a one standard deviation higher "non-genetic" risk aversion is associated with a 4.4% lower chance of holding equity. For three of the four variables, the signs of the genetic and non-genetic portions are identical and across the eight regressions, and the genetic portion, on average, accounts for 42% of the regression R^2 (i.e., the average value in the bottom row across columns 1, 2, 4, 6, 7, and 9 is 42%.

Similar to Panel A, the regression of orthogonalized market participation on the chance the market increases by more than 20% is unique—as the coefficients associated with the genetic and non-genetic portions have opposite signs. Specifically, the genetic component is negative suggesting it may capture the genetic share of investor views of the riskiness of stock returns, while the non-genetic component is positive suggesting it may better capture investor views of very favorable returns when excluding risk aversion and the other beliefs from the analysis.

We next examine the genetic and non-genetic components when including genetic and non-genetic components of all four measures (risk aversion and the three belief metrics) as regressors. Recognize, however, that because the genetic components are fitted values from the same eight explanatory variables (i.e., the PGSs), variation in the genetic components are highly correlated. Specifically, the absolute value of the correlation between the four predicted (i.e., genetic) components averages 66% versus 15% for the four residual (non-genetic) components.[51] As a result, the four

[51] For example, Panel B in Table III shows that the genetic endowment associated with Neuroticism is positively associated with risk aversion, beliefs regarding the likelihood of at least 20% gain in equity prices, and beliefs regarding the likelihood of at least a 20% decline in equity prices. As a result, an increase in the Neuroticism PGS will be associated with a simultaneous increase in the genetic components of risk aversion, beliefs regarding the likelihood of at least 20% increase in equity markets, and beliefs regarding the likelihood of at least a 20% decline in equity markets. Recognize that Table III values are based on regressions of risk aversion or beliefs whereas the genetic components used in Table IV are based on estimates of risk aversion or beliefs once first accounting for variation in risk aversion and beliefs attributed to the control variables and the 10 genetic principal components. As shown in the Internet Appendix, however, the relations between outcomes and the PGSs (Table III) are nearly identical to the relation between orthogonalized outcomes and the PGSs that are used in Table IV.

genetic components suffer from high levels of collinearity reducing the power of the tests. Importantly, however, the collinearity does not bias the R^2 (or how the R^2 is apportioned between the genetic and non-genetic components).

Results in the fifth column of Panel B suggest that the eight genetic endowments account for approximately 30% of the relation between equity market participation and risk aversion and beliefs regarding the distribution of equity returns. Similarly, results in the last column suggest that the eight genetic endowments can explain about 20% of the relation between the fraction of assets held in equity and risk aversion and beliefs.

IV. Genetics, Explanatory Variables used in Previous Studies, and Outcomes

We next estimate the extent to which the eight genetic endowments help link the relations between outcome variables (risk aversion, beliefs regarding the distribution of expected equity returns, and market participation) and the 11 variables documented by previous work to help explain equity market participation. We begin by evaluating the relations between the outcome variables and each of the 11 individual explanatory variables. Because many of the 11 explanatory variables are strongly correlated (e.g., as shown in Table II Panel C, the correlation between optimism and trust is 0.49) and our goal is understanding the role of the eight genetic endowments in explaining these previously identified relations (rather than determining which of the 11 explanatory variables best explains the outcome variables), we focus on regressions of the outcome variables on the control variables (indicators for HRS waves, gender, age, married, and retired) and each of the 11 explanatory variables individually. In addition, for ease of interpretation and comparison, we standardize (i.e., rescale to zero mean, unit variance) each of the 11 explanatory variables. The results of the 66 panel regressions (11 explanatory variables times six outcomes) are reported in Panel A of Table V.

[Insert Table V about here]

The first column reports that each of the 11 explanatory variables used in previous work to explain equity market participation are related to risk aversion in the expected direction consistent with the explanation that these variables correlate with equity market participation, at least in part, because they are associated with variation in risk aversion. Specifically, greater wealth, income, education, cognition, trust, sociability, optimism, height, and health are all associated with a greater willingness to take risk. Conversely, early life poverty (*poor*) and BMI are positively are associated with risk aversion.

The next three columns report the relations between the 11 explanatory variables and beliefs regarding the distribution of expected returns. This is largely new territory—few studies examine if

these explanatory variables are related to variation in investor beliefs regarding the likelihood of a positive return, and, as far as we are aware, no published study examines whether these variables may be associated with variation in beliefs regarding the shape of the distribution of expected returns (i.e., the likelihood of a large gain (>20%) or loss (<-20%) in equity valuations). The results in the second column reveal that 10 of the 11 explanatory variables are related to beliefs regarding of the likelihood of a market increase consistent with the hypothesis that variables correlate with equity market participation, at least in part, because they reflect heterogeneity in beliefs. Specifically, wealth, income, education, cognition, trust, sociability, optimism, and health are all associated with a higher expected probability of the market increasing in the next year, whereas early life poverty and higher BMI are associated with beliefs of a lower likelihood of a positive market return in the next year.

The results for the next two columns provide important and intuitive insights as well. First, as noted above, because individuals greatly underestimate the likelihood of the stock market rising, we expect lower probabilities of at least a 20% gain associated with more realistic views of the standard deviation of equity returns. Consistent with our hypothesis, cognition and height are negatively associated with beliefs regarding the likelihood of a greater than 20% equity market return. Trust, optimism, and health, however, are positively related to variation in perceptions regarding the likelihood of a greater than 20% gain in equity markets.

Recall that, on average, individuals greatly overestimate the likelihood of a greater than 20% decline in equity markets. Thus, we expect variables associated with a better understanding of equities, such as cognition or education, to be negatively associated with beliefs regarding the likelihood of a greater than 20% decline, while the "negative" variables (BMI and early life poverty) should be positively associated with beliefs regarding the likelihood of a greater than 20% decline. The results in the fourth column confirm this pattern. Specifically, all 11 variables have the expected sign and 10 of the 11 are statistically significant at the 5% level or better.

Consistent with previous work, equity market participation (last two columns) is positively related to wealth, income, education, cognition, trust, sociability, optimism, height, and health and negatively related to poverty in early life and BMI.

In sum, the results in the first four columns suggest these 11 explanatory variables are related to equity market participation, at least in part, because they are associated with both heterogeneity in risk aversion and heterogeneity in beliefs regarding the distribution of expected equity returns. For instance, individuals with high trust are less risk averse (first column), their beliefs regarding the

distribution of expected equity returns is shifted to the right (columns 2, 3, and 4), and they are more likely to participate in equity markets (last two columns),

Because all independent variables are standardized in both Tables III and V, we can directly examine the relative magnitude of the relation between the genetic endowments (established at least 50 years prior to the outcomes) and the outcome variables versus the relation between the explanatory variables (measured approximately simultaneously with the outcomes) used in previous studies and the outcome variables. For example, the average absolute value of the coefficient in the first column of Panel A in Table V reveals that a one standard deviation change in the average explanatory variable used in previous work is associated with a 12.5% standard deviation change in risk aversion (recall that risk aversion is also standardized). That is, the average absolute value of coefficients in the first column of Table V Panel A is 0.125. The average absolute value in the first column of Table III Panel A is 0.039. Thus, on average, the relation between risk aversion and the average genetic endowment (measured more than 50 years prior) is about one-third the relation between risk aversion and the average contemporaneously measured explanatory variable used in previous studies (i.e., 0.039/0.125=0.31). Results are similar across the two equity participation variables (i.e., the average absolute value in the last two columns of Table III are 30.6% and 30.4%, respectively, of the average absolute value of the last two columns in Panel A of Table V). The relations for the three measures of beliefs regarding the distribution of expected equity returns and the average genetic endowment range from 25.4% to 97.9% of the relation between the beliefs regarding the distribution of expected equity returns and the average of the 11 explanatory variables. Note that we do not include the genetic endowments and the explanatory variables used in previous studies as independent variables in the same regression because the 11 explanatory variables are realizations while the PGSs are predictors, e.g., the Height PGS (established at conception) helps predict the respondent's height, but the respondent's height is the realized value of the phenotype when risk aversion, beliefs, and stock market participation are measured.

Panel B in Table V examines the relation between the outcome measures and the 11 explanatory variables when including all 11 explanatory variables in the model simultaneously (and the controls for gender, HRS wave, age, retired, and married). Our results are not directly comparable to previous work because: (1) different combinations of explanatory variables (e.g., because it had yet to be "discovered," Hong, Kubik, and Stein (2004) do not attempt to include trust in their tests), and (2) many of our metrics, although measuring the same construct, differ from early measures. As noted above, for instance, advances in the HRS survey allow for, arguably, better measures of variables such

as trust and optimism. Regardless, the results in Panel B of Table V largely support previous work. For example, the fifth column reveals that equity market participation is positively related to wealth, income, education, cognition, trust, and health, and is inversely related to early life poverty. Analogously, risk aversion (first column of Panel B) is negatively related to income, education, sociability, optimism, height, and health. Inconsistent with the tests in Panel A, however, risk aversion is positively related to trust when including the other 10 explanatory variables in the model. Moreover, the variables are largely associated with the beliefs regarding the distribution of expected stock returns in the expected direction (with the exception of a positive relation between BMI and beliefs regarding the likelihood of a market increase). For example, lower wealth, lower trust, lower health, and early life poverty are all associated with beliefs that a severe (greater than 20%) market decline in the next year is more likely.

A. Genetic Endowments and Explanatory Variables used in Previous Studies

Having shown the consistency of our work with the previous studies, we now provide original insights by examining whether these commonly used 11 explanatory variables are related to the eight genetic endowments. Specifically, we estimate panel regressions of each of the 11 explanatory variables on the control variables (indicator variables for HRS waves, gender, age, married, and retired), the 10 HRS genetic principal components, and each of the individual PGSs. For ease of interpretation, we standardize (rescale to zero mean, unit variance) each of the 11 dependent variables. Panel A in Table VI reports the results. Thus, for example, the top left-hand cell implies that a one standard deviation higher Educational Attainment PGS is associated with a 13.4% standard deviation higher log wealth.

[Insert Table VI about here]

The results in Table VI provide strong evidence that all 11 explanatory variables identified in previous work are meaningfully related to a number of the eight genetic endowments. Moreover the signs of the relations are intuitive. Wealth, for example, is positively related to the Educational Attainment PGS and Height PGS and inversely related to the Neuroticism PGS, Depressive Symptoms PGS, Myocardial Infarction PGS, Coronary Artery Disease PGS, and BMI PGS. In fact, every explanatory variable is meaningfully related to at least three of the eight PGSs and the average explanatory variable is related to 5.4 of the eight PGSs.

Panel B of Table VI repeats the analysis but regresses each of the 11 explanatory variables on all eight PGS (and the control variables and 10 HRS genetic principal components) simultaneously. Each of the 11 explanatory variable is meaningfully related to at least two of the PGSs in the multiple

regression framework. Once again, the signs of the relations are largely intuitive. Optimism, for instance, is positively related to the Educational Attainment PGS but negatively related to the PGSs associated with Neuroticism, Depressive Symptoms, Myocardial Infarction, and BMI. In sum, the results in Table VI reveal strong evidence that variation in all 11 of the traditional variables we examine is driven, at least in part, by the eight genetic endowments we evaluate.

B. Genetic Endowments, Explanatory Variables used in Previous Studies, and Outcomes

We next estimate the role of the eight genetic endowments in explaining the relation between the explanatory variables used in previous studies and the outcome variables. Following our previous analysis, we begin by removing variation in the outcome variables and the explanatory variables used in previous studies related to the control variables and the genetic principal components. Specifically, we regress each of the outcome variables (as in Table IV) and each of the explanatory variables on the control variables (indicators for HRS waves, age, gender, retired, and married) and the first 10 principal components of the genetics data. The residuals from these regressions—denoted orthogonalized outcomes or orthogonalized explanatory variables—reflect variation in each outcome or explanatory variable that cannot be explained by the control variables or the 10 genetic principal components.

We then regress each of the orthogonalized explanatory variables onto the eight PGSs and denote the portion of the orthogonalized explanatory variable predicted by the eight PGSs (i.e., the fitted value) as the genetic component of the explanatory variable and the portion unexplained by the eight PGSs (i.e., the residual) as the non-genetic component (the Internet Appendix reports the coefficient estimates from these regressions). As before, for ease of interpretation, we standardize (rescale to unit variance and zero mean) both the genetic and non-genetic components and then regress each of the orthogonalized outcome variables on genetic and non-genetic components of each of the orthogonalized explanatory variable. Because the genetic and non-genetic components are mechanically independent, analogous to Equations (2) and (3), the R^2 from these regressions can be directly partitioned into the portion attributed to the genetic versus non-genetic components. Note that because these regressions include only two mechanically independent standardized variables (i.e., the genetic portion and the non-genetic portion), the fraction of the R^2 attributed to the genetic portion is simply the ratio of the squared value of the genetic coefficient to the sum of the squared values of the genetic and non-genetic components.

Panel A in Table VII reports pairs of coefficients from the 66 regressions (six outcomes times 11 explanatory variables).[52] Panel B reports the R^2 from the regression of the orthogonalized outcome variable on the portion of the orthogonalized explanatory variable predicted by the eight PGSs and the portion of the orthogonalized explanatory variable not explained by the eight PGSs. Panel C reports the portion of the regression R^2 attributed to the "genetic" portion of the explanatory variable.

The results in Table VII suggest that a meaningful portion of the relation between the each of the 11 explanatory variables used in previous studies and the outcomes arises from the portion of the explanatory variable associated with the eight genetic endowments. For example, in a regression of orthogonalized risk aversion on the portion of orthogonalized log wealth associated with PGSs (i.e., the standardized fitted value from a regression of orthogonalized log wealth on the eight PGSs) and the portion of orthogonalized log wealth independent of PGSs (i.e., the standardized residual value from the regression of orthogonalized log wealth on the eight PGSs), a one standard deviation higher genetic component of log wealth is associated with 5.3% standard deviation lower risk aversion (recall risk aversion is standardized). A one standard deviation larger "non-genetic" log wealth is associated with an 11.7% standard deviation lower risk aversion. Therefore, as shown in the top cell of Panel C, we estimate that approximately 17% of the relation ($0.053^2/(0.053^2+0.117^2)$) between orthogonalized risk aversion and the two components of orthogonalized wealth arises from the portion of orthogonalized wealth associated with the eight genetic endowments and 83% of the relation arises from the portion of orthogonalized wealth independent of the genetic endowments.[53]

As shown in the first column of the bottom row in Table VII, averaged across the 11 explanatory variables, 20% of the ability of the these explanatory variables to explain orthogonalized risk aversion arises from the portion of the variable associated with the eight genetic endowments. The bottom row of Panel C reports averages ranging from 20% to 58% across the six outcome variables.

For eight of the 11 explanatory variables—income, education, cognition, trust, sociability, early life poverty, BMI, and health—the portion associated with the eight genetic endowments explains, on average, at least 33% of the variable's explanatory share across the six outcome variables (i.e., the average value in the Panel C row associated with that explanatory variable is at least 33%). In short,

[52] Because the genetic component of each of these 11 components are estimated from the same eight PGSs, we cannot include all 11 genetic components simultaneously as regressors (analogous to the fifth and tenth columns in Table IV), i.e., the matrix is mechanically singular once including more than eight of the 11 genetic components.

[53] Similar to Table IV, the signs of the coefficients in "PGS" column of Panel A sometimes differ from the signs of the "Non-PGS" column only in the case of orthogonalized beliefs regarding the likelihood of a greater than 20% market return.

the results in Table VII reveal that much of the explanatory power from variables shown to explain equity market participation is associated with the eight genetic endowments.

V. Conclusions

Genetic endowments associated with cognition, personality, health, and body shape, established at least 50 years prior, are associated with equity market participation and the two factors that traditional economic theory suggests play the central role in determining equity market participation—risk aversion and beliefs regarding the distribution of expected equity returns. These relations are intuitive and consistent with our hypotheses—risk aversion is inversely related to genetic endowments associated with Educational Attainment and Height and positively related to genetic endowments associated with Neuroticism, Depressive Symptoms, Myocardial Infarction, and Coronary Disease. Similarly, beliefs regarding the distribution of expected equity returns are associated with genetic endowments. For instance, the genetic endowment associated with Neuroticism is negatively related to beliefs regarding the likelihood of the market rising next year, but positively related to beliefs regarding the likelihood of extreme market movements (up or down by more than 20%). Consistent with the patterns in risk aversion and beliefs, equity market participation is positively related to genetic endowments associated with Educational Attainment, General Cognition, and Height and negatively related to genetic endowments associated with Neuroticism, Depressive Symptoms, Myocardial Infarction, Coronary Disease, and BMI. In short, the nitrogenous bases in our DNA—for example, whether we have an A or a G at a specific locus—predict our risk aversion, our beliefs regarding the distribution of equity returns, and our stock market participation.

Consistent with traditional economic theory, equity market participation is positively related to positive beliefs (i.e., higher expected return and lower risk) regarding the distribution of expected equity returns and negatively related to risk aversion. When we partition beliefs and risk aversion into the component explained by the genetic endowments and the portion orthogonal to the genetic endowments, the portion associated with the eight genetic endowments accounts for about 30% of the relation between equity market participation and the two factors that drive the decision in traditional economic theory: risk aversion and beliefs regarding the distribution of expected equity returns.

We also find that 11 explanatory variables used in previous work are related to equity market participation consistent with this research. Moreover, consistent with economic theory and the hypothesis that these variables are related to equity market participation, at least in part, because they

are related to heterogeneity in risk aversion and heterogeneity in beliefs regarding the distribution of expected equity returns, we find that all 11 explanatory variables are related to both risk aversion and beliefs regarding the distribution of equity returns. For instance, beliefs regarding the likelihood of a greater than 20% decline in equity prices are negatively related to wealth, income, education, cognition, trust, sociability, optimism and health, and positively related to negative early life economic experiences and BMI.

Each of the 11 explanatory variables documented in previous work are also associated with the genetic endowments we examine. Trust, for example, is positively related to genetic endowments associated with Educational Attainment and General Cognition, but negatively related to genetic endowments associated with Neuroticism, Depressive Symptoms, Myocardial Infarction, and BMI. When we partition the relation between the outcome variables (risk aversion, beliefs regarding the distribution of expected equity returns, and equity market participation), and the 11 explanatory variables into the portion of the explanatory variables related to the genetic endowments and the portion of the explanatory variables orthogonal to the genetic endowments, we estimate that 20%-58% of the relation is attributed to the portion associated with the eight genetic endowments.

Our results have implications for understanding heterogeneity in investor risk aversion, beliefs regarding the distribution of expected equity returns, and, as a result, heterogeneity in equity market participation. The statistically significant relations between the outcome variables and genetic endowments imply that all four genetic components we examine (cognition, personality, health, and body shape) meaningfully predict outcomes more than half a century later. For instance, for two otherwise identical individuals, the person with high genetic risk for Neuroticism is likely to be more risk averse, believe there is a greater likelihood of a market crash in the next year, and be less likely to hold equity.

The genetic endowments we examine also help in understanding both why explanatory variables documented in previous work are associated with equity market participation, and why individuals differ in these characteristics. For instance, for two otherwise identical individuals, the person born with a higher Educational Attainment PGS is likely to have greater optimism, and greater optimism is associated with lower risk aversion and beliefs of a higher likelihood of a positive market return—all of which can help explain the positive relation between optimism and equity market participation.

REFERENCES

Abraham, Gad, Aki S. Havulinna, Oneil G. Bhalala, Sean G. Byars, Alysha M. De Livera, Laxman Yetukuri, …, and Michael Inouye, 2016, Genomic prediction of coronary heart disease, *European Heart Journal* 37, 3267-3278.

Abraham, Gad, Adam Kowalczyk, Justin Zobel, and Michael Inouye, 2013, Performance and robustness of penalized and unpenalized methods for genetic prediction of complex human disease, *Genetic Epidemiology* 37, 184-195.

Abraham, Gad, and Michael Inouye, 2014, Fast principal component analysis of large-scale genome-wide data, *PLoS ONE* 9, e93766.

Addoum, Jawad M., George Korniotis, and Alok Kumar, 2017, Stature, obesity, and portfolio choice, *Management Science* 63, 3393-3413.

Anagol, Santosh, Vimal Balasubramaniam, and Tarun Ramadorai, 2018, Noise trading and experience effects, working paper, University of Pennsylvania, University of Warwick, and Imperial College London.

Ashraf, Quamrul, and Oded Galor, 2013, The 'Out of Africa' hypothesis, human genetics, and comparative economic development, *American Economic Review* 103, 1-46.

Balloch, Adnan, Anamaria Nicolae, and Dennis Philip, 2015, Stock market literacy, trust, and participation, *Review of Finance* 19, 1925-1963.

Barnea, Amir, Henrik Cronqvist, and Stephan Siegel, 2010, Nature or nurture: What determines investor behavior? *Journal of Financial Economics* 98, 583-604.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom, 2018, Genetic endowments and wealth inequality, NBER working paper no. 24642.

Beauchamp, Jonathan P., David Cesarini, Magnus Johannesson, Matthijs J. H. M. van der Loos, Phillipp D. Koellinger, Patrick J. F. Groenen, …, and Nicholas A. Christakis, 2011, Molecular genetics and economics, *Journal of Economic Perspectives* 25, 57-82.

Benartzi, Shlomo, 2001, Excessive extrapolation and the allocation to 401(k) accounts to company stock, *Journal of Finance* 56, 1747-1764.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, 2013, Who is "behavioral"? Cognitive ability and anomalous preferences, *Journal of the European Economic Association* 11, 1231-1255.

Benjamin, Daniel J., David Cesarini, Christopher F. Chabris, Edward L. Glaeser, David I. Laibson, Vilmundur Gudnason, …, and Paul Lichtenstein, 2012, The promises and pitfalls of genoeconomics, *Annual Review of Economics* 4, 627-662.

Berkman, Lisa F., Ichiro Kawachi, and M. Maria Glymour, 2014, *Social Epidemiology* (Oxford University Press, New York).

Bernile, Gennaro, Vineet Bhagwat, and P. Raghavendra Rau, 2017, What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior, *Journal of Finance* 70, 167-206.

Bharath, Sreedhar T., and DuckKi Cho, 2016, Ephemeral experiences, long lived impact: Disasters and portfolio choice, Working paper, Arizona State University.

Bonaparte, Yosef, and Alok Kumar, 2013, Political activism, information costs, and stock market participation, *Journal of Financial Economics* 107, 760-786.

Brennan, Thomas J., and Andrew W. Lo, 2011, The origins of behavior, *Quarterly Journal of Finance* 1, 55-108.

Brennan, Thomas J., and Andrew W. Lo, 2012, An evolutionary model of bounded rationality and intelligence, *PLoS ONE* 7, e50310.

Brennan, Thomas J., Andrew W. Lo, and Ruixun Zhang, 2018, Variety is the spice of life: Irrational behavior as adaptation to stochastic environments, *Quarterly Journal of Finance* 8, 1-39.

Brown, Sue, and Richard Sias, 2019, The fault in our stars: Molecular genetics and technology adoption, Working paper, University of Arizona.

Bricker, Jesse, and Geng Li, 2017, Credit scores, social capital, and stock market participation, Finance and Economics Discussion Series Working Paper No. 2017-008.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott Weisbenner, 2008, Neighbors matter: Casual community effects and stock market participation, *Journal of Finance* 63, 1509-1531.

Bush, William S. and Jason H. Moore, 2012, Genome-wide association studies, *PLoS Computational Biology* 8, e1002822.

Calvet, Laurent E. and Paolo Sodini, 2014, Twin picks: Disentangling the determinants of risk-taking in household portfolios, *Journal of Finance* 69, 867-906.

Cameron, Lisa, and Manisha Shah, 2015, Risk-taking behavior in the wake of natural disasters, *Journal of Human Resources* 50, 484-515.

Campbell, John Y., 2006, Household finance, *Journal of Finance* 61, 1533-1604.

Caplin, Andrew, and John Leahy, 2001, Psychological expected utility theory and anticipatory feelings, *Quarterly Journal of Economics* 116, 55-79.

CARDIoGRAMplusC4D Consortium, 2015, A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease, *Nature Genetics* 47, 1121-1130.

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, 809-842.

Cesarini, David, Magnus Johannesson, Patrik K. E. Magnusson, and Björn Wallace, 2012, The behavioral genetics of behavioral anomalies, *Management Science* 58, 21-31.

Cesarini, David, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace, 2009, Heritability of overconfidence, *Journal of the European Economic Association* 7, 617-627.

Cesarini, David, Magnus Johannesson, Paul Lichtenstein, Örjan Sandewall, and Björn Wallace, 2010, Genetic variation in financial decision making, *Journal of Finance* 65, 1725-1754.

Cesarini, David, and Peter M. Visscher, 2016, Genetics and educational attainment, *npj Science of Learning* 2, 1-7.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, 2009, Reinforcement learning and savings behavior, *Journal of Finance* 64, 2515-2534.

Christelis, Dimitris, Tullio Jappelli, and Mario Padula, 2010, Cognitive abilities and portfolio choice, *European Economic Review* 54, 18-38.

Cohn, A., J. Engelmann, E. Fehr, and M. Marechal, 2015, Evidence for counter-cyclical risk aversion: An experiment with financial professionals, *American Economic Review* 105, 860-885.

Cole, Shawn A., Anna Paulson, and Gauri Kartini Shastry, 2014, Smart money? The effect of education on financial outcomes, *Review of Financial Studies* 27, 2022-2051.

Cotti, Chad, Richard A. Dunn, and Nathan Tefft, 2015, The Dow is killing me: Risky health behaviors and the stock market, *Health Economics* 24, 803-821.

Cronqvist, Henrik, and Stephan Siegel, 2014, The genetics of investment biases, *Journal of Financial Economics* 113, 215-234.

Cronqvist, Henrik, and Stephan Siegel, 2015, The origins of savings behavior, *Journal of Political Economy* 123, 123-169.

Curcuru, Stephanie, John Heaton, Deborah Lucas, and Damien Moore, 2010, Heterogeneity and portfolio choice: Theory and evidence, in *Handbook of Financial Econometrics* (Elsieiver, Amsterdam), 337-382.

Davies, G., N. Armstrong, J. C. Bis, J. Bressler, V. Chouraki, S. Giddaluru, … and I. J. Deary, 2015, Genetic contributions to variation in general cognitive function: A meta-analysis of genome-wide association studies in the CHARGE consortium (N=53,949), *Molecular Psychiatry* 20, 183-192.

de Moor, M. H., S. M. van den Berg, K. J. Verweij, R. F. Krueger, M. Luciano, … and D. I. Boomsma, 2015, Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder, 2015, *JAMA Psychiatry* 72, 642-650.

Deaton, Angus and Raksha Arora, 2009, Life at the top: The benefits of height, *Economics and Human Biology* 7, 133-136.

Decker, Simon and Hendrik Schmitz, 2016, Health shocks and risk aversion, *Journal of Health Economics* 50, 156-170.

Dick, Danielle M., 2011, Gene-environment interaction in psychological traits and disorders, *Annual Review of Clinical Psychology* 7, 383-409.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, and Kim Peijnenburg, 2016, Ambiguity aversion and household portfolio choice puzzles: Empirical evidence, *Journal of Financial Economics* 119, 559-577.

Dohman, Thomas, Armin Falk, David Huffman, and Uwe Sunde, 2010, Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100, 1238-1260.

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, 369-379.
Dominitz, Jeff, and Charles F. Manski, 2011, Measuring and interpreting expectations of equity returns, *Journal of Applied Econometrics* 26, 352-370.
Dow, James, and Sérgio Ribeiro da Costa Werlang, 1992, Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica* 60, 197-204.
Engelberg, Joseph, and Christopher A. Parsons, 2016, Worrying about the stock market: Evidence from hospital admissions, *Journal of Finance* 71, 1227-1250.
Evangelou, Evangelos, and John P.A. Ioannidis, 2013, Meta-analysis methods for genome-wide association studies and beyond, *Nature Reviews Genetics* 14, 379-389.
Evans, Anthony M., and William Revelle, 2008, Survey and behavioral measurements of interpersonal trust, *Journal of Research in Personality* 42, 1585-1593.
Fischhoff, Baruch, and Wändi Bruine De Bruin, 1999, Fifty–fifty=50%? *Journal of Behavioral Decision Making* 12, 149-163.
Fisher, Gwenith G., Halimah Hassan, Jessica D. Faul, Willard L. Rodgers, and David R. Weir, 2017, Health and Retirement Study imputation of cognitive functioning measures: 1992-2014 (Final Release Version), available at http://hrsonline.isr.umich.edu/modules/meta/xyear/cogimp/desc/COGIMPdd.pdf.
Georgarakos, Dimitris, and Giacomo Pasini, 2011, Trust, sociability, and stock market participation, *Review of Finance* 15, 693-725.
Giannetti, Mariassunta, and Tracy Yue Wang, 2016, Corporate scandals and household stock market participation, *Journal of Finance* 71, 2591-2636.
Girirajan, Santhosh, 2017, Missing heritability and where to find it, *Genome Biology* 18, 89.
Goetzmann, William N., Dasol Kim, and Robert J. Shiller, 2017, Crash beliefs from investor surveys, NBER working paper no. 22143.
Goldstein, Benjamin A., Lingyao Yang, Elias Salfati, Themistoclies L. Assimes, 2015, Contemporary considerations for constructing a genetic risk score: An empirical approach, *Genetic Epidemiology* 39, 439-445.
Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa, 2011, IQ and stock market participation, *Journal of Finance* 66, 2121-2164.
Guillemette, Michael A., Michael Finke, and John Gilliam, 2012, Risk tolerance questions to best determine client portfolio allocation preferences, *Journal of Financial Planning* 5, 36-44.
Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2008, Trusting the stock market, *Journal of Finance* 63, 2557-2600.
Guo, Guang, 2005, Twin studies: What can they tell us about nature and nurture? Contexts 4, 43-47.
Harlow, W. V. and Keith C. Brown, 1990, Understanding and assessing financial risk tolerance: A biological perspective, *Financial Analysts Journal* 46, 50-80.
Heaton, John and Deborah Lucas, 1997, Market frictions, savings behavior, and portfolio choice, *Macroeconomic Dynamics* 1, 76-101.
Heimer, Rawley Z., 2014, Friends do let friends buy stocks actively, *Journal of Economic Behavior & Organization*, 107, 527-540.
Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social interaction and stock-market participation, *Journal of Finance* 59, 137-163.
Hurd, Michael D., and Susann Rohwedder, 2012, Stock price expectations and stock trading, NBER working paper no. 17973.
Hurd, Michael, Maarten Van Rooij, and Joachim Winter, 2011, Stock market expectations of Dutch households, *Journal of Applied Econometrics* 26, 416-436.
Joseph, Jay, 2013, The use of classical twin method in the social and behavioral sciences: The fallacy continues, *Journal of Mind and Behavior* 34, 1-39.
Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi, 2012, A careful re-examination of seasonality in international sock markets: Comment on sentiment and stock returns, *Journal of Banking and Finance* 36, 934-956.

Katti, Gunnar, Lars Olov Bygren, and Soren Edvinsson, 2002, Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period, *European Journal of Human Genetics* 10, 682-688.

Kaustia, Markku, and Samuli Knupfer, 2008, Do investors overweight personal experience? Evidence from IPO subscriptions, *Journal of Finance* 63, 2679-2702.

Kaustia, Markku, and Sami Torstila, 2011, Stock market aversion? Political preferences and stock market participation, *Journal of Financial Economics* 100, 98-112.

Keller, Evelyn Fox, and David Harel, 2007, Beyond the gene, *PLoS ONE* 2, e1231.

Kézdi, Gábor, and Robert J. Willis, 2003, Who becomes a stockholder? Expectations, subjective uncertainty, and asset allocation, unpublished working paper, Central European University and University of Michigan.

Klein, Robert J., Caroline Zeiss, Emily Y. Chew, Jen-Yue Tsai, Richard S. Sackler, Chad Haynes, …, and Josephine Hoh, 2005, Complement factor H polymorphism in age-related macular degeneration, *Science* 308, 385-389.

Knupfer, Samuli, Elias Rantapuska, and Matti Sarvimaki, 2017, Formative experiences and portfolio choices: Evidence from the Finnish great depression, *Journal of Finance* 72, 133-166.

Lee, Boram, Leonard Rosenthal, Chris Veld, and Yulia Veld-Merkoulova, 2015, Stock market expectations and risk aversion of individual investors, *International Review of Financial Analysis* 40, 122-131.

Lillard, Lee, and Robert Willis, 2001, Cognition and wealth: The importance of probabilistic thinking, unpublished working paper, University of Michigan.

Locke, A. E., B. Kahali, S. I. Berndt, A. E. Justice, T. H. Pers, F. R. Day, … and D. C. Croteau-chonka, 2015, Genetic studies of body mass index yield new insights for obesity biology, *Nature* 518, 197-206.

Luppino, Floriana S., Leonore M. de Wit, Paul F. Bouvy, Theo Stijnen, Pim Cuijpers, Brenda W. J. H. Penninx, and Frans G. Zitman, 2010, Overweight, obesity, and depression—A systematic review and meta-analysis of longitudinal studies, *Archives of General Psychiatry* 67, 220-229.

Malmendier, Ulrike, Geoffrey Tate, and Jon Yan, 2011, Overconfidence and early life experiences: The effect of managerial traits on corporate financial policies, *Journal of Finance* 66, 1687-1733.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics* 126, 373-416.

Malmendier, Ulrike, and Stefan Nagel, 2016, Learning from inflation experiences, *Quarterly Journal of Economics* 131, 53-87.

Martin, Alicia R., Christopher R. Gignoux, Raymond K. Walters, Genevieve L. Wojcik, Benjamin M. Neale, Simon Gravel, …, and Eimear E. Kenny, 2017, Human demographic history impacts genetic risk prediction across diverse populations, *The American Journal of Human Genetics* 100, 635-649.

McInerney, Melissa, Jennifer M. Mellor, and Lauren Hersch Nicholas, 2013, Recession depression: Mental health effects of the 2008 stock market crash, *Journal of Health Economics* 32, 1090-1104.

Meeks, Suzanne and Stanley A. Murrell, 2001, Contribution of education to health and life satisfaction in older adults mediated by negative affect, *Journal of Aging and Health* 13, 92-119.

Merton, Robert C., 1969, Lifetime portfolio selection under uncertainty: The continuous-time case, *Review of Economics and Statistics* 51, 247-257.

Merton, Robert C., 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373-413.

Nicholson, Nigel, Emma Soane, Mark Fenton-O'Creevy, and Paul Willman, 2005, Personality and domain-specific risk taking, *Journal of Risk Research* 8, 157-176.

Obeidat, Ma'en, Louise V. Wain, Nick Shrine, Noor Kalsheker, Maria Soler Artigas, Emmanouela Repapi, …, and Veronique Vitart, 2011, A comprehensive evaluation of potential lung function associated genes in the SpiroMeta general population sample, *PLoS ONE* 6, e19382.

Okbay, Aysu, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, and Tune H. Pers, 2016, Genome-wide association study identifies 74 loci associated with educational attainment, *Nature* 533, 539-542.

Pasco, Julie A, Lana J. Williams, Felice N. Jacka, Sharon L. Brennan, and Michael Berk, 2013, Obesity and the relationship with positive and negative affect, *Australian and New Zealand Journal of Psychiatry* 45, 477-482.

Persico, Nicola, Andrew Postelwaite, and Dan Silverman, 2004, The effect of adolescent experience on labor market outcomes: The case of height, *Journal of Political Economy* 112, 1019-1053.

Polderman, Tinca J.C., Beben Benyamin, Christiaan A de Leeuw, Patrick F Sullivan, Arjen van Bochoven, Peter M Visscher, …, and Danielle Posthuma, 2015, Meta-analysis of the heritability of human traits based on fifty years of twin studies, *Nature Genetics* 47, 702-709.

Puri, Manju, and David T. Robinson, 2007, Optimism and economic choice, *Journal of Financial Economics* 86, 71-99.

Rosen, Harvey S., and Stephen Wu, 2004, Portfolio choice and health status, *Journal of Financial Economics* 72, 457-484.

Sacerdote, Bruce, 2011, Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? Chapter 5, J. Benhabib, M. Jackson & A. Bisin (Eds.), Handbook of Social Economics, Amsterdam, North Holland.

Schunkert, H., I. R. Konig, S. Kathiresan, M. P. Reilly, T. L. Assimes, H. Holm, … and D. Absher, 2011, Large scale association analysis identifies 13 new susceptibility loci for coronary artery disease, *Nature Genetics* 43, 333-338.

Shin, Su Hyun, Dean R. Lillard, and Jay Bhattacharya, 2018, Understanding the correlation between Alzheimer's disease polygenic risk, wealth, and the composition of wealth holdings, working paper, University of Alabama, Ohio State University, DIW-Berlin, NBER, and Stanford University.

Simonson, Matthew A., Amanda G. Wills, Matthew C. Keller, and Matthew B. McQueen, 2011, Recent methods for polygenic analysis of genome-wide data implicate an important effect of common variants on cardiovascular disease risk, *BMC Medical Genetics* 12, 146.

Smith, Jacqui, Lindsay Ryan, Gwenith G. Fisher, Amanda Sonnega, and David Weir, 2017, Psychosocial and lifestyle questionnaire 2006-2016, Documentation Report, Core Section LB, available at: https://hrs.isr.umich.edu/publications/biblio/9066.

Stoolmiller, Michael, 1999, Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior-genetic adoption studies, *Psychological Bulletin* 115, 392-409.

Torvik, Fartein Ask, Audun Welander-Vatn, Eivind Ystrom, Gun Peggy Knudsen, Nikolai Czajkowski, Kenneth S. Kendler, and Ted Richborn-Kjennerud, 2016, Longitudinal associations between social anxiety disorder and avoidant personality disorder: A twin study, *Journal of Abnormal Psychology* 125, 114-124.

Turkheimer, Eric, 2000, Three laws of behavior genetics and what they mean, *Current Directions in Psychological Science* 5, 160-164.

Vineis, Paolo, and Neil E. Pearce, 2011, Genome-wide association studies may be misinterpreted: Genes versus heritability, *Carcinogenesis* 32, 1295-1298.

Visscher, Peter M., Naomi R. Wray, Qian Zhang, Pamela Sklar, Mark I. McCarthy, Matthew A. Brown, ..., and Jian Yang, 2017, 10 years of GWAS discovery: Biology, function, and translation, *The American Journal of Human Genetics* 101, 5-22.

Vissing-Jorgensen, Annette, 2002, Towards an explanation of household portfolio choice heterogeneity: Nonfinanical income and participation cost structures, NBER working paper no. 8884.

Ware, Erin B., Lauren L. Schmitz, Jessica Faul, Arianna Gard, Colter Mitchell, Jennifer A. Smith, … Sharon L.R. Kardia, 2017, Heterogeneity in polygenic scores for common human traits, *bioRxiv*, 106062.

Ware, Erin, Lauren Schmitz, Arianna Gard, and Jessica Faul, 2018, HRS documentation report: HRS polygenic scores – Release 2, Survey Research Center, University of Michigan.

Winerman, Lea, 2004, A second look at twin studies, *Monitor on Psychology* 35, 46.

Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, … and N. Amin, 2014, Defining the role of common variation in the genomic and biological architecture of adult human height, *Nature Genetics* 46, 1173-1186.

Yogo, Motohiro, 2016, Portfolio choice in retirement: Health risk and the demand for annuities, housing and risky assets, *Journal of Monetary Economics* 80, 2016.

Zhang, Ruixun, Brennan, Thomas J., and Andrew W. Lo, 2014a, Group selection as behavioral adaptation to systematic risk, *PLoS ONE* 9, e110848.

Zhang, Ruixun, Brennan, Thomas J., and Andrew W. Lo, 2014b, The origins of risk aversion, *Proceedings of the National Academy of Sciences* 111, 17777-17782.

Zuk, Or, Eliana Hechter, Shamil R. Sunyaev, and Eric S. Lander, 2012, The mystery of missing heritability: Genetic interactions create phantom heritability, *Proceedings of the National Academy of Sciences* 109, 1193-1198.

Appendix A – Variable Detail and Construction

Panel A: HRS computed polygenic scores (Source: HRS Polygenic Scores – Release 2, 2006-2012 Genetic Data, April 2018)	
Educational Attainment PGS	Weights from the 2016 Social Science Genetic Association Consortium (SSGAC) based on 293,723 individuals in the discovery sample and 111,349 individuals in the replication sample (see Okbay et al. (2016))
General Cognition PGS	Weights from the 2015 Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) based on 53,949 individuals who took multiple diverse cognition tests (see Davies et al. (2015))
Neuroticism PGS	Weights from the 2016 Social Science Genetic Association Consortium (SSGAC) based on 170,911 individuals (see Okbay et al. (2016) and de Moor et al. (2015))
Depressive Symp. PGS	Weights from the 2016 Social Science Genetic Association Consortium (SSGAC) based on 180,866 individuals (see Schunker et al. (2011))
Myocardial Infarc. PGS	Weights from the 2015 Coronary ARtery DIsease Genome wide Replication And Meta-analysis (CARDIoGRAM) consortium based on 184,305 individuals (see CARDIoGRAMplusC4D Consortium (2015))
Coronary Disease PGS	Weights from the 2011 Coronary ARtery DIsease Genome wide Replication And Meta-analysis (CARDIoGRAM) consortium based on 86,995 individuals (see Okbay et al. (2016) and de Moor et al. (2015))
BMI PGS	Weights from the 2015 Genetic Investigation of ANthropometric Traits (GIANT) consortium based on samples totaling 322,154 individuals (see Locke et al. (2015))
Height PGS	Weights from the 2014 Genetic Investigation of ANthropometric Traits (GIANT) consortium based on 253,288 individuals in the discovery sample and 80,067 individuals in the replication sample (see Wood et al. (2014))
Panel B: Economic outcome variables (Source: Combination of RAND HRS fat files and HRS raw data files)	
Risk aversion	In the 2014 and 2016 waves, respondents were asked "Are you generally a person who tries to avoid taking risks or one who is fully prepared to take risks? Please rate yourself from 0 to 10, where 0 means 'not at all willing to take risks' and 10 means 'very willing to take risks.'" We subtract the respondent's average (over years 2014 and 2016) score to this question from 10 and then rescale the value to zero mean and unit variance (for ease of interpretation).
$P(R_m > 0)$	Respondent response in each HRS wave to the question, "We are interested in how well you think the economy will do in the future. 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?"
$P(R_m > 20\%)$	Respondent response in each HRS wave to the question, "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 have gained in value by more than 20 percent compared to what they are worth today?"
$P(R_m < -20\%)$	Respondent response in each HRS wave to the question, "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 have fallen in value by more than 20 percent compared to what they are worth today?"
Equity participation	Equity participation equals one if the respondent reports holding equities in (i) their pension fund(s), (ii) IRA/KEOGH accounts, or (iii) directly. (i) Pension fund information comes from the employment and pension section of the interview. Specifically, respondents are asked how much money is in the pension plan now and what percent of this plan is invested in stock? This question is asked of both spouses for households with partners (i.e., the data are at the respondent level). If either spouse reports any of their pension wealth (respondents are asked about multiple pensions if they have more than one pension) is invested in the stock market we code the household (associated with the financial respondent) as participating in equity markets via their pension assets. In 2010, the respondent is limited to a maximum of four pension plans (i.e., eight per household for partnered households). In 2012, HRS changed the structure of the pension section and no longer limited the number of pension funds to four per respondent. We find, however, that very few respondents report more than four pension plans. Respondents who answer they are not sure what fraction is invested in equities, are asked a series of "unfolding" brackets to approximate the amount, e.g., is it less than 40% and more than 20%? In cases where the respondent answer is greater than zero, we code the household as investing in the stock market. (ii) The IRA/KEOGH information comes from the financial respondent's interview regarding household assets and income. Specifically, the financial respondent is asked, "Do you [or your] [husband/wife/partner] currently have any money or assets that are held in an Individual Retirement Account, that is, in an IRA or KEOGH account?" For those households that respond positively, the financial respondent is asked if any of that IRA/KEOGH is invested in stocks or mutual funds, how much is in each account, and what fraction of the IRA/KEOGH is invested in stock. If the household has more than one IRA/KEOGH, the financial respondent is asked about them in order of size up to the three largest IRA/KEOGH accounts. If the financial respondent reports any of the household

	IRA/KEOGH accounts are at least partially invested in the stock market we code the household as participating in their IRA/KEOGH. (iii) The direct holdings information also arises from the assets and income section of the HRS interview. Specifically, financial respondents are asked "Aside from anything you have already told me about, do you [or your] [husband/wife/partner] have any shares of stock or stock mutual funds?" Respondents who answer yes are classified as participating directly (i.e., in non-retirement accounts) in equity markets following previous work (e.g., Addoum, Korniotis, and Kumar (2017), Hong, Kubik, and Stein (2004)). In addition, respondents who answer yes to this question are also asked the value of the stock holdings as well as the value other financial assets.
%Equity	We compute the fraction of all (including retirement) wealth invested in equities as sum of the value of direct stocks holdings, IRA/KEOGHs held in stock (inferred from IRA/KEOGH account values and percent invested in stock), and pension funds held in stock (inferred from the pension fund account values and percent invested in stock) divided by the sum of total financial wealth, IRA/KEOGH account values, and pension fund account values. We follow RAND (see https://www.rand.org/labor/aging/dataprod.html) and compute non-housing financial wealth as the sum of checking/savings/money market funds, CDs/government savings bonds/Treasury bills, corporate/municipal/government or foreign bonds/bond funds, other savings or assets, less other (non-housing) debt. We compute the fraction of direct financial wealth invested in equities as the value of direct equity holdings (see iii above) divided by non-housing financial wealth.
Panel C: Explanatory variables used in previous studies (Source: Unless otherwise noted, combination of RAND HRS longitudinal file (v2), HRS 2014 tracker file, RAND HRS fat files, HRS raw files)	
Wealth	Positive hypothesized relation with equity market participation due to lower direct and indirect participation costs including, for example, "acquiring enough information about risks and returns to determine the household's optimal mix between stocks and riskless assets" (Vissing-Jorgensen (2002)). Given most individuals greatly underestimate the mean return and overestimate market risk, learning should shift the expected return beliefs right and shrink the market risk estimate (i.e., probabilities associated with extreme returns). Moreover, if relative risk aversion is decreasing then wealth will be inversely related to risk aversion (see, for instance, Calvet and Sodini (2014)). Thus, wealth may impact participation via direct costs, risk aversion, and beliefs.

	We measure raw real (CPI-adjusted to 2010 dollars) wealth defined as the sum of financial wealth (using the RAND financial wealth definition detailed above), net (loan-adjusted) value of homes, net (loan-adjusted) value of other real estate, value of automobiles, value of business, and IRA/KEOGH value. We winsorize raw real wealth at the 1% and 99% level. The minimum raw real winsorized wealth is slightly greater than -$66,000. Therefore we add $66,000 to our real wealth variable (to ensure households with negative wealth are not excluded from our data) and take the natural logarithm to mitigate skewness.
Income	Positive hypothesized relation with equity market participation due to lower direct and indirect participation costs (e.g., Vissing-Jorgensen (2002)). Directly analogous to wealth, higher incomes may be associated with lower direct costs, lower risk aversion, and more optimistic beliefs regarding the distribution of expected returns. We also use the RAND definition of total household income which is the sum of household labor earnings (wages, professional fees, bonus, income from second job), self-employment income, income from investments (real estate, business, stocks, bonds, checking/savings, CDs, other income), income from pensions, social security income, unemployment/workers compensation income, governmental transfers (welfare, veteran income, food stamps), and any other income. We convert all values to 2010 dollars and winsorize at the 1% and 99% level before taking the natural logarithm.
Education	Positive hypothesized relation with equity market participation due to improved understanding of financial markets and cognitive ability. Education may also impact participation via higher labor income. See Cole, Paulson, and Shastry (2014) for evidence of the positive relation between stock market participation and education. Empirical evidence suggests cognitive ability is inversely related to risk aversion (e.g., Dohmen, Falk, Huffman, and Sunde (2010)). As noted above, given the typical individual underestimates the mean return and overestimates market risk, learning (associated with education) should shift the expected return beliefs right and shrink the probabilities associated with extreme returns. We use years of education as our measure of education.
Cognitive ability	Positive hypothesized relation with equity market participation due to improved information processing skill, income, wealth, & education. As noted for education (above), evidence suggests cognitive ability is inversely related to risk aversion and greater cognitive ability should shift beliefs regarding the likelihood of the market rising higher and the likelihood of an extreme market return lower. For evidence regarding cognitive ability and market participation see Kezdi and Willis (2003), Benjamin, Brown, Shapiro (2013), Christelis, Jappelli, and Padula (2010), Grinblatt, Keloharju, and Linnainmaa (2011), and Cole, Paulson and Shasty (2014).

	Following Fisher, Hassan, Faul, Rodgers, and Weir (2017), we use the HRS total cognition measure which ranges from 0-35 based on seven different tests. The first two tests are word recall—in the immediate recall tests, respondents are given a list of 10 nouns (based on four possible lists) and asked to immediately recall as many as possible (score 0-10). The second test is delayed recall. After about five minutes of other questions, respondents are asked a second time to recall as many words as possible (score 0-10). The third test is serial 7s—respondents are asked to subtract 7 from 100 and continue to do so for five trials. Score is the number of correct subtractions (score 0-5). The fourth test is backwards counting—respondents are given two chances to count backwards for 10 consecutive numbers starting from both 20 and 86. The recorded score is 2 points if the task is performed correctly on the first try, 1 point if completed correctly on the second try, and zero if not correct on either attempt. The fifth test asks respondent to name the date and day of week—one point for correct day, year, month, and day of week (0-4 points). The sixth test asks respondents to name two objects "What do you usually use to cut paper?" and "What do you call the kind of prickly plant that grows in the desert?" Scores are 1 point for each correct answer (0-2 points). The seventh test asks respondents to name the current President and Vice-President. One point for each correct answer (0-2 points). To ensure a larger sample, we use the average cognition score by each respondent over any of the four waves (2010, 2012, 2014, 2016) where the respondent completed the cognition tests. The 2010-2014 cognition scores include values imputed by HRS (See http://hrsonline.isr.umich.edu/modules/meta/xyear/cogimp/desc/COGIMPdd.pdf). The 2016 score is based on raw data and does not include imputed values. (Source: HRS Cognition dataset for 2010-2014, HRS raw cognition data for 2016)
Trust	Positive hypothesized relation with equity market participation due to lower subjective probability of being cheated by equity markets (e.g., Guiso, Sapienza, and Zingales (2008), Giannetti and Wang (2016)). Perceived likelihood of being cheated by markets results in lower expected return and a higher probability of an extreme left tail return (see Balloch, Nicolae, and Philip (2015)) for individuals with lower trust. To measure trust, the Psychosocial and Lifestyle Questionnaire asks respondents to rate their agreement with five statements, where score 1=strongly disagree, 2=somewhat disagree, 3=slightly disagree, 4=slightly agree, 5=somewhat agree, 6=strongly agree. The five statements are: (1) Most people dislike putting themselves out to help other people, (2) Most people will use somewhat unfair means to gain profit or an advantage rather than lose it, (3) No one cares much what happens to you, (4) I think most people would lie in order to get ahead, and (5) I commonly wonder what hidden reasons another person may have for doing something nice for me. Following the guidance in Smith, Ryan, Fisher, Sonnega, and Weir (2017), we compute the average score for individuals who rate at least three of the five statements.

	The trust metric is included in the Leave-Behind Questionnaires for 2006, 2008, 2010, and 2012. To ensure a larger sample, we use the average trust score by each respondent over any of the four waves (2006, 2008, 2010, 2012) where the respondent completed the trust questions. Because higher values in the raw data indicate lower trust, we compute trust as 7 less the respondent average trust score over all waves in which they participate from 2006-2012. Thus, trust ranges from 1 to 6 with high values indicating greater trust. Note that the psychosocial and lifestyle questionnaire measures "cynical hostility," a term from psychology to describe the "routine lack of trust of other people" (see Berkman, Kawachi, and Glymour (2014)). For consistency with the finance literature, we refer to this dimension as "trust."
Sociability	Positive hypothesized relation with equity market participation due to (1) more informed regarding equities via word-of-mouth and observational learning, and (2) utility from socializing about investing (e.g., Hong, Kubik, and Stein (2004), Guiso, Sapienza and Zingales (2008), Heimer (2014)). As noted above, because most individuals underestimate the mean return and overestimate market risk, learning (via socialization) should shift the expected return beliefs right and shrink the probabilities associated with extreme returns. Hong Kubik, and Stein (2004) propose that more social individuals may be more "bold" (less risk averse), but find no evidence of a meaningful relation between their measures of sociability and risk tolerance. Evidence outside of finance, however, reports a negative relation between risk aversion and sociability (e.g., Nicholson, Soane, Fenton-O'Creevy, and Willman (2005)). The HRS Psychosocial and Lifestyle Questionnaire (Leave-Behind Questionnaire) asks respondents three questions: On average, how often do you do each of the following with any of your friends, not counting any who live with you? (1) meet up? (2) speak on the phone? (3) write or email? Respondents answers are (1) three or more times a week, (2) once or twice a week, (3) once or twice a month, (4) every few months, (5) once or twice a year, or (6) less than once a year or never. Following the guidance in Smith, Ryan, Fisher, Sonnega, and Weir (2017), we score sociability as the average score for respondents who respond to at least two of the three questions. Because the raw scores range from 1-6 and higher values indicate less sociability, we subtract the value from 6 so that sociability ranges from 1 to 6 with higher values implying greater sociability. Because Leave-Behind Questionnaires are given to half the respondents in each wave, we use the respondents' average sociability score over waves they participate in from 2006-2016.
Optimism	Positive hypothesized relation with equity market participation due to higher expected outcomes (e.g., Puri and Robinson (2007)). Empirically, Puri and Robinson find an inverse relation between risk aversion and optimism. Thus, higher optimism may be associated with lower risk aversion, beliefs of a higher likelihood of markets rising, and beliefs of a lower likelihood of extreme left tail returns.

	Individual optimism measures come from the Psychosocial and Lifestyle Questionnaire. Specifically, respondents are asked to rate their agreement with six statements, where score 1=strongly disagree, 2=somewhat disagree, 3=slightly disagree, 4=slightly agree, 5=somewhat agree, 6=strongly agree. The six statements are: (1) If something can go wrong for me it will, (2) I'm always optimistic about my future, (3) In uncertain times, I usually expect the best, (4) Overall, I expect more good things to happen to me than bad, (5) I hardly ever expect things to go my way, and (6) I rarely count on good things happening to me. Following Smith, Ryan, Fisher, Sonnega, and Weir (2017), we compute the raw optimism score by multiplying items (1), (5), and (6) by -1 and then averaging across all six values for respondents who answer at least four of the statements. We then compute the respondent average raw optimism score by computing the average over waves they participate in between 2006-2016. Because this value can (theoretically) range from -6 to +6, we add seven so trust is a scale from 1 to 13 where higher values indicate greater trust.
Negative early life experience	Negative hypothesized relation with equity market participation due to increased risk aversion and/or lower expected returns (e.g., Malmendier & Nagel (2011)). The authors also find evidence consistent with the expected returns explanation but point out that such evidence does not preclude the risk aversion channel. We capture whether individuals grew up with negative early life economic experiences by a dummy variable if any of the four following conditions were met: (1) Respondents answer "poor" to the question, "Now think about your family when you were growing up, from birth to age 16. Would you say your family during that time was pretty well off financially, about average, or poor?", (2) Respondent answers "yes" to the question, "While you were growing up, before age 16, did financial difficulties ever cause you or your family to move to a different place?", (3) Respondent answers "yes" to the question, "Before age 16, was there a time when you or your family received help from relatives because of financial difficulties?", and (4) Respondent answers "yes" to the question, "Before age 16, was there a time of several months or more when your father had no job?"
Height	Positive hypothesized relation with equity market participation due to lower risk aversion. Addoum, Korniotis, and Kumar (2017) propose that the impact of height on risk aversion may also operate via educational attainment, sociability, and positive early life experience. Dohmen, Falk, Huffman, and Sunde (2010) report empirical evidence that height is inversely related to risk aversion. Persico, Postlewaite, and Silverman (2004) find that the positive relation between height and labor income is primarily accounted for by teen (rather than adult) height suggesting that social effects (such as greater self-esteem) drive the height-wage correlation. Work (see Deaton and Arora (2009)) also suggests that

	height is positively associated with greater positive emotions (e.g., enjoyment) and fewer negative emotions (e.g., sadness). As detailed above, such social/optimism effects could manifest in higher expectations for market returns. Following Addoum, Korniotis, and Kumar (2017), we measure relative height (which we denote height) as the respondent's height in meters less the average height of individuals of the same age and gender in the same wave.
BMI	Negative hypothesized relation with equity market participation due to the positive relation between BMI and risk aversion. Addoum, Korniotis, and Kumar (2017) propose the relation between BMI, risk aversion, and stock market participation may also operative via educational attainment, sociability, and early life experience. Empirically, there is a positive association between BMI and depression (see meta-analysis by Luppino et al. (2010)). Work (e.g., Pasco, Williams, Jacka, Brennan, and Berg (2013)) also demonstrates that BMI is positively related to negative affect (e.g., distress, anger, disgust, fear, and shame) which may result in more pessimistic view of future stock returns. Following Addoum, Korniotis, and Kumar (2017), relative BMI (computed in each wave as respondent's weight in kilograms divided by the respondents height in meters squared) is computed as respondent's BMI less the average BMI for individuals of the same age and gender in the same wave.
Health	Rosen and Wu (2004) suggest health may impact portfolio choice due to changes in "the marginal utility of consumption, degree of risk aversion, rate of time preference, and variability of income." Although the authors find a strong relation between health and equity participation, they find little evidence regarding the channel linking portfolio decisions to health. Other work demonstrates, however, that health is strongly inversely related to negative affect (e.g., see Table 2 in Meeks and Murrell (2001)) and that negative health shocks are associated with increased risk aversion (e.g., Decker and Schmitz (2016)). Therefore, we hypothesize that health will be inversely related to risk aversion, positively related to more optimistic beliefs regarding the distribution of equity returns, and (consistent with Rosen and Wu's evidence), positively related to equity market participation. Health from each respondent's response is determined from the question "Would you say your health is: excellent, very good, good, fair, or poor?" Responses are coded on a five point scale where excellent=5 and poor=1 (such that higher values reflect better health).
Panel D: Control variables (Source: Combination of RAND HRS longitudinal file (v2), HRS 2014 tracker file, RAND HRS fat files, HRS raw files))	
Age indicators	Indicator variables for respondent's age at time of interview (31 indicators for age 50-80).

HRS wave indicators	Indicator variables for HRS waves (2010, 2012, 2014, and 2016).
Gender indicator	Indicator variable equals one if respondent is male.
Retired indicator	Indicator variable equals one if respondent's labor force status is retired.
Married indicator	Indicator variable equals one if respondent is married at time of interview.

Table I
Descriptive Statistics

This table reports summary statistics for the Polygenic scores (Panel A), the outcome variables (Panel B), 11 explanatory variables used in previous studies (Panel C), and control variables (Panel D). The sample period includes HRS waves 2010, 2012, 2014, and 2016. Appendix A provides details for all variables.

	N	Mean	Standard Deviation	Minimum	Maximum
Panel A: Polygenic score (PGS) variables					
Educational Attainment PGS	12,633	0.0	1.0	-3.745	3.462
General Cognition PGS	12,633	0.0	1.0	-3.921	3.956
Neuroticism PGS	12,633	0.0	1.0	-3.744	3.482
Depressive Symptoms PGS	12,633	0.0	1.0	-3.365	3.609
Myocardial Infarction PGS	12,633	0.0	1.0	-3.641	3.205
Coronary Artery Disease PGS	12,633	0.0	1.0	-3.718	3.398
BMI PGS	12,633	0.0	1.0	-3.646	4.095
Height PGS	12,633	0.0	1.0	-4.355	2.672
Panel B: Economic outcome variables					
Raw risk aversion	12,633	4.109	2.038	0.0	10.0
Risk aversion	12,633	0.0	1.0	-2.016	2.891
$P(R_m > 0)$	12,633	47.344	26.525	0.0	100.0
$P(R_m > 20\%)$	12,633	27.644	22.658	0.0	100.0
$P(R_m < -20\%)$	12,633	30.274	23.198	0.0	100.0
Equity participation	12,633	0.624	0.485	0.0	1.0
%Equity	10,234	0.381	0.385	0.0	1.488
Panel C: Explanatory variables used in previous studies					
ln(Wealth)	12,633	12.508	1.195	7.801	15.356
ln(Income)	12,633	10.694	1.176	5.656	13.261
Years education	12,633	13.874	2.382	0.0	17.0
Cognition	11,119	24.111	3.620	7.667	34.0
Trust	11,601	4.178	0.987	1.0	6.0
Sociability	11,780	3.949	0.922	1.0	6.0
Optimism	12,097	8.069	0.908	3.800	9.622
Grew up poor	12,630	0.400	0.490	0.0	1.0
Height	12,632	0.011	0.067	-0.336	0.364
BMI	12,547	-0.201	5.952	-17.168	34.866
Health	12,631	3.394	1.006	1.0	5.0
Panel D: Control variables					
Male	12,633	0.487	0.500	0.0	1.0
Age	12,633	66.282	8.239	50.0	80.0
Retired	12,633	0.478	0.500	0.0	1.0
Married	12,633	0.610	0.488	0.0	1.0

Table II
Correlations

This table reports correlation coefficients for the Polygenic scores (Panel A), the outcome variables (Panel B), and 11 explanatory variables used in previous studies (Panel C). The sample period includes HRS waves 2010, 2012, 2014, and 2016. Appendix A provides details for all variables.

Panel A: Polygenic score (PGS) variables

	Edu. Attain.	Gen. Cognition	Neuroticism	Depressive	Myocardial	Coronary	BMI	Height
Edu. Attain.	1.000							
Gen. Cognition	0.269	1.000						
Neuroticism	-0.065	0.054	1.000					
Depressive	-0.092	-0.007	0.548	1.000				
Myocardial	-0.142	-0.076	0.015	0.046	1.000			
Coronary	-0.115	-0.085	0.042	0.046	0.414	1.000		
BMI	-0.161	-0.060	-0.127	-0.025	0.087	0.025	1.000	
Height	-0.092	-0.091	0.009	-0.044	0.076	-0.007	-0.173	1.000

Panel B: Economic outcome variables

	Risk aversion	$P(R_m > 0)$	$P(R_m > 20\%)$	$P(R_m < -20\%)$	Equity participation	%Equity
Risk aversion	1.000					
$P(R_m > 0)$	-0.125	1.000				
$P(R_m > 20\%)$	-0.014	0.510	1.000			
$P(R_m < -20\%)$	0.038	-0.240	-0.006	1.000		
Equity participation	-0.141	0.221	0.026	-0.100	1.000	
%Equity	-0.130	0.201	0.030	-0.089	0.624	1.000

Panel C: Explanatory variables used in previous studies

	ln(Wealth)	ln(Income)	Education	Cognition	Trust	Sociability	Optimism	Poor	Height	BMI	Health
ln(Wealth)	1.000										
ln(Income)	0.371	1.000									
Education	0.320	0.315	1.000								
Cognition	0.201	0.285	0.431	1.000							
Trust	0.189	0.090	0.222	0.191	1.000						
Sociability	0.159	0.106	0.231	0.212	0.242	1.000					
Optimism	0.267	0.188	0.253	0.249	0.487	0.233	1.000				
Poor	-0.091	-0.099	-0.156	-0.096	-0.080	-0.041	-0.068	1.000			
Height	0.059	0.081	0.099	0.090	0.020	0.062	0.074	-0.038	1.000		
BMI	-0.173	-0.053	-0.077	-0.031	-0.097	-0.026	-0.118	0.054	-0.062	1.000	
Health	0.255	0.230	0.232	0.255	0.202	0.146	0.359	-0.125	0.045	-0.240	1.000

Table III
Regression of Equity Market Participation, Risk Aversion, and Beliefs on Genetic Endowments

We regress each of outcome variables on control variables (indicators for HRS waves, age, gender, retired, and married), the first 10 principal components of the genetics data, and each of the eight genetic PGSs individually and report results in Panel A. Panel B reports regression coefficients when including all eight PGSs as regressors simultaneously (and the control variables and 10 principal components). All PGSs are standardized (rescaled to zero mean unit variance). Appendix A provides details for all variables. In all cases, standard errors are clustered at the respondent level. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	Risk aversion	$P(R_m > 0)$	$P(R_m > 20\%)$	$P(R_m < -20\%)$	Equity participation	%Equity
Panel A: Outcomes on individual PGSs (+control variables and 10 principal components)						
Edu. Attainment PGS	-0.041***	1.829***	-0.475*	-0.767***	0.065***	0.039***
Gen. Cognition PGS	-0.025	0.652**	-0.629**	-0.817***	0.022***	0.017***
Neuroticism PGS	0.075***	-0.927**	0.773**	1.201***	-0.038***	-0.017***
Depressive Symp. PGS	0.030**	-0.427	0.322	0.653**	-0.030***	-0.020***
Myocardial Infarc. PGS	0.033**	-1.010***	-0.153	-0.096	-0.021***	-0.010*
Coronary Disease PGS	0.049***	0.004	0.217	-0.184	-0.011*	-0.008
BMI PGS	0.011	-0.399	0.189	0.323	-0.033***	-0.019***
Height PGS	-0.050*	0.817	-0.302	-0.514	0.025**	0.001
Panel B: Outcomes on all eight PGSs (+control variables and 10 principal components)						
Edu. Attainment PGS	-0.022	1.644***	-0.255	-0.468*	0.054***	0.033***
Gen. Cognition PGS	-0.014	0.260	-0.527*	-0.662**	0.007	0.009
Neuroticism PGS	0.072***	-0.640	0.731**	1.021***	-0.020**	-0.002
Depressive Symp. PGS	-0.006	0.135	-0.029	0.143	-0.012*	-0.014**
Myocardial Infarc. PGS	0.011	-1.063***	-0.354	-0.137	-0.011*	-0.003
Coronary Disease PGS	0.040**	0.635*	0.289	-0.246	0.001	-0.003
BMI PGS	0.003	0.003	0.136	0.215	-0.021***	-0.012**
Height PGS	-0.037	0.414	-0.169	-0.318	0.011	-0.007
R^2	4.69%	5.15%	1.34%	3.30%	10.18%	6.20%
Number of clusters	5,560	5,560	5,560	5,560	5,560	4,831
Number of obs.	12,633	12,633	12,633	12,633	12,633	10,234

Table IV
Equity Market Participation and the Genetic Components of Risk Aversion and Beliefs

Each of the outcome variables (risk aversion, beliefs regarding the distribution of equity returns, and equity participation measures) is regressed on the control variables (indicator variables for HRS waves, gender, age, married, and retired) and the first 10 principal components of the genetics data. We define the residuals from these regressions as the orthogonalized outcome variables. Panel A reports regressions of orthogonalized equity market participation on orthogonalized risk-aversion and orthogonalized beliefs. Panel B reports coefficients from regressions of orthogonalized equity market participation on the portion of orthogonalized risk-aversion and orthogonalized beliefs explained by the eight PGSs and the portion of orthogonalized risk aversion and beliefs unrelated to the eight PGSs. The bottom row in Panel B reports the portion of the R^2 attributed to genetic variation in orthogonalized risk aversion and beliefs. Standard errors are clustered at the respondent level. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	Equity participation					%Equity				
Panel A: Equity market participation on risk aversion and beliefs regarding expected return distribution										
Risk aversion	-0.048***				-0.039***	-0.037***				-0.032***
$P(R_m > 0)$		0.088***			0.098***		0.064***			0.068***
$P(R_m > 20\%)$			0.018***		-0.035***			0.015***		-0.021***
$P(R_m < -20\%)$				-0.046***	-0.023***				-0.033***	-0.016***
R^2	1.06%	3.58%	0.15%	0.99%	5.04%	1.00%	2.94%	0.15%	0.78%	4.14%
Panel B: Equity market participation on portion of risk aversion and beliefs related to PGSs and portion orthogonal to PGSs										
Risk aversion PGS	-0.049***				-0.002	-0.029***				0.001
$P(R_m > 0)$ PGS		0.061***			0.052***		0.031***			0.024***
$P(R_m > 20\%)$ PGS			-0.045***		-0.023			-0.030***		-0.023***
$P(R_m < -20\%)$ PGS				-0.053***	0.002				-0.026***	0.004
Risk aversion non-PGS	-0.044***				-0.036***	-0.035***				-0.030***
$P(R_m > 0)$ non-PGS		0.083***			0.091***		0.062***			0.065***
$P(R_m > 20\%)$ non-PGS			0.020***		-0.031***			0.016***		-0.018***
$P(R_m < -20\%)$ non-PGS				-0.043***	-0.022***				-0.032***	-0.016***
R^2	2.03%	4.96%	1.11%	2.17%	6.36%	1.48%	3.41%	0.82%	1.19%	4.70%
$\%R^2$ due to PGS	55.72%	35.25%	84.13%	59.80%	30.44%	39.81%	20.09%	77.95%	39.34%	19.50%

Table V
Regression of Outcomes on Explanatory Variables Used in Previous Studies

Panel A reports regressions of the outcome variables (risk aversion, three measures of beliefs regarding the distribution of expected returns, and two measures of equity market participation) on the control variables (indicator variables for HRS waves, gender, age, married, and retired) and each of the 11 explanatory variables individually (i.e., 66 regressions for 11 explanatory variables times six outcomes). Panel B reports results from regressions of the outcome variables on the control variables and the 11 explanatory variables simultaneously. All independent variables are standardized (i.e., rescaled to zero mean, unit variance). Appendix A provides details for all variables. In all cases, standard errors are clustered at the respondent level. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	Risk aversion	$P(R_m > 0)$	$P(R_m > 20\%)$	$P(R_m < -20\%)$	Equity participation	%Equity
Panel A: Outcomes on controls and explanatory variables individually						
ln(Wealth)	-0.132***	4.859***	0.360	-2.514***	0.217***	0.121***
ln(Income)	-0.117***	3.358***	-0.129	-1.361***	0.135***	0.080***
Education	-0.170***	4.693***	-0.317	-1.086***	0.139***	0.081***
Cognition	-0.134***	4.327***	-0.808***	-1.081***	0.123***	0.070***
Trust	-0.106***	4.065***	0.596**	-1.787***	0.093***	0.053***
Sociability	-0.174***	2.390***	-0.127	-0.664**	0.074***	0.042***
Optimism	-0.224***	3.879***	0.551**	-1.289***	0.092***	0.053***
Poor	0.033**	-0.887**	-0.141	0.732***	-0.051***	-0.021***
Height	-0.064***	0.458	-0.586**	-0.021	0.016***	0.008*
BMI	0.055***	-0.488*	0.208	0.579**	-0.048***	-0.017***
Health	-0.173***	3.384***	0.477*	-1.646***	0.110***	0.047***
Panel B: Outcomes on controls and all 11 explanatory variables simultaneously						
ln(Wealth)	-0.022	2.698***	0.450	-1.977***	0.176***	0.101***
ln(Income)	-0.038***	0.707**	-0.149	-0.003	0.025***	0.022***
Education	-0.092***	2.084***	-0.062	0.401	0.043***	0.026***
Cognition	-0.010	1.461***	-1.377***	-0.479	0.036***	0.021***
Trust	0.048**	1.865***	0.703**	-1.385***	0.030***	0.018***
Sociability	-0.085***	0.194	-0.091	0.071	0.006	0.007
Optimism	-0.153***	1.000**	0.396	0.267	-0.001	0.009
Poor	-0.017	0.035	-0.137	0.498*	-0.016***	-0.002
Height	-0.037**	-0.169	-0.497*	0.211	-0.006	-0.001
BMI	0.014	0.837***	0.498*	-0.290	0.000	0.008
Health	-0.095***	1.428***	0.815***	-0.668**	0.032***	0.009
R^2	12.35%	10.73%	1.29%	3.75%	31.00%	15.93%
Number of clusters	4,357	4,357	4,357	4,357	4,357	3,827
Number of obs.	9,222	9,222	9,222	9,222	9,222	8,151

Table VI
Regression of Explanatory Variables used in Previous Studies on Polygenic Scores (PGSs)

Panel A reports regression coefficients from regressions of each of the 11 explanatory variables on the control variables (indicator variables for HRS waves, gender, age, married, and retired), the first 10 principal components of the genetics data and each of the eight PGSs individually. Panel B reports regression coefficients when including all eight PGSs as regressors simultaneously. Both dependent and independent variables are standardized (rescaled to zero mean, unit variance). Appendix A provides details for all variables. In all cases, standard errors are clustered at the respondent level. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	ln(Wealth)	ln(Inc.)	Education	Cognition	Trust	Social	Optimism	Poor	Height	BMI	Health
Panel A: Explanatory variables on individual PGS (+10 principal components)											
Edu. Attainment PGS	0.134***	0.092***	0.228***	0.177***	0.136***	0.072***	0.129***	-0.044***	0.046***	-0.059***	0.103***
Gen. Cognition PGS	0.024	0.019	0.096***	0.102***	0.053***	-0.011	0.018	0.006	0.008	-0.015	0.038***
Neuroticism PGS	-0.057***	-0.011	-0.082***	-0.094***	-0.116***	-0.067***	-0.127***	0.058***	0.012	-0.014	-0.073***
Depressive Symp. PGS	-0.052***	-0.011	-0.055***	-0.051***	-0.084***	-0.065***	-0.106***	0.035**	-0.004	-0.001	-0.065***
Myocardial Infarc. PGS	-0.049***	-0.033***	-0.062***	-0.054***	-0.037**	-0.005	-0.048***	0.023	-0.042***	0.039***	-0.069***
Coronary Disease PGS	-0.035***	-0.029***	-0.043***	-0.012	-0.005	-0.021	-0.012	0.012	-0.012	0.007	-0.043***
BMI PGS	-0.075***	-0.046***	-0.058***	-0.058***	-0.074***	-0.006	-0.064***	0.042***	-0.014	0.277***	-0.092***
Height PGS	0.049*	0.013	0.078***	0.068**	0.021	0.041	0.062**	-0.052*	0.701***	-0.002	0.033
Panel B: Explanatory variables on all eight PGS (+10 principal components)											
Edu. Attainment PGS	0.111***	0.079***	0.211***	0.148***	0.103***	0.055***	0.109***	-0.023	0.021	0.000	0.063***
Gen. Cognition PGS	-0.019	-0.011	0.036**	0.056***	0.018	-0.021	-0.013	0.018	-0.035**	0.002	0.010
Neuroticism PGS	-0.033	-0.012	-0.041*	-0.060***	-0.077***	-0.037	-0.071***	0.040*	0.030	0.005	-0.036*
Depressive Symp. PGS	-0.018	0.009	0.001	0.005	-0.033*	-0.039*	-0.055***	-0.005	0.011	-0.031*	-0.024
Myocardial Infarc. PGS	-0.021	-0.005	-0.027	-0.041**	-0.032*	0.003	-0.045**	0.024	-0.021	0.014	-0.046***
Coronary Disease PGS	-0.013	-0.022*	-0.006	0.027	0.032*	0.001	0.007	-0.016	0.015	-0.019	-0.006
BMI PGS	-0.060***	-0.023*	-0.015	-0.021	-0.049***	-0.004	-0.043**	0.027	-0.006	0.282***	-0.064***
Height PGS	0.008	0.000	0.000	0.027	-0.012	0.016	0.044	-0.062**	0.714***	0.028	0.014
R^2	12.86%	23.26%	12.80%	15.51%	7.08%	4.53%	6.77%	1.90%	18.22%	8.33%	6.14%
Number of clusters	4,357	4,357	4,357	4,357	4,357	4,357	4,357	4,357	4,357	4,357	4,357
Number of obs.	9,922	9,922	9,922	9,922	9,922	9,922	9,922	9,922	9,922	9,922	9,922

Table VII
Outcomes and the Genetic Components of Explanatory Variables used in Previous Studies

Each of the outcome variables (risk aversion, beliefs regarding the distribution of equity returns, and equity participation measures) and each of the traditional explanatory variables (e.g., wealth, income, education, etc.) are regressed on the control variables (indicator variables for HRS waves, gender, age, married, and retired) and the first 10 principal components of the genetics data. We define the residuals from these regressions as the orthogonalized outcome variables and the orthogonalized traditional explanatory variables. Each orthogonalized outcome variable is regressed on the portion of orthogonalized traditional explanatory variable explained by the eight PGSs and the portion of the orthogonalized traditional explanatory variable unrelated to the eight PGSs. Panel B reports the R^2 from each regression and Panel C reports the portion of the R^2 attributed to genetic variation in the traditional explanatory variable. Appendix A provides details for all variables. Standard errors are clustered at the respondent level. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table VII (continued)
Outcomes and the Genetic Components of Explanatory Variables used in Previous Studies

	Risk aversion		$P(R_m > 0)$		$P(R_m > 20\%)$		$P(R_m < -20\%)$		Equity participation		%Equity	
Panel A: Outcomes on portion of traditional explanatory variable related to PGSs and portion orthogonal to PGSs												
	PGS	Non-PGS	PGS	Non-PGS	PGS	Non-PGS	PGS	Non-PGS	PGS	Non-PGS	PGS	Non-PGS
ln(Wealth)	-0.053***	-0.117***	1.741***	4.325***	-0.476*	0.382	-0.798***	-2.226***	0.068***	0.196***	0.039***	0.099***
ln(Income)	-0.048***	-0.097***	1.685***	2.728***	-0.411	-0.121	-0.629**	-1.082***	0.064***	0.112***	0.038***	0.062***
Education	-0.054***	-0.155***	1.875***	4.178***	-0.558**	-0.227	-0.896***	-0.818***	0.066***	0.122***	0.038***	0.070***
Cognition	-0.064***	-0.113***	1.897***	3.763***	-0.575**	-0.671**	-1.013***	-0.833***	0.064***	0.105***	0.038***	0.056***
Trust	-0.058***	-0.094***	1.867***	3.728***	-0.537**	0.690**	-1.094***	-1.592***	0.068***	0.081***	0.037***	0.045***
Sociability	-0.051***	-0.167***	1.281***	2.194***	-0.442*	-0.141	-0.723***	-0.546**	0.057***	0.067***	0.030***	0.038***
Optimism	-0.052***	-0.213***	1.684***	3.630***	-0.470*	0.667***	-0.971***	-1.156***	0.065***	0.081***	0.034***	0.046***
Poor	0.056***	0.029**	-1.515***	-0.747**	0.460*	-0.152	0.835***	0.650**	-0.062***	-0.045***	-0.029***	-0.019***
Height	-0.025**	-0.057***	0.885***	0.385	-0.103	-0.430*	-0.249	-0.074	0.021***	0.014**	0.007	0.008
BMI	0.012	0.051***	-0.466	-0.326	0.192	0.183	0.317	0.491*	-0.033***	-0.040***	-0.019***	-0.012**
Health	-0.057***	-0.161***	1.586***	3.080***	-0.489*	0.530**	-0.813***	-1.495***	0.065***	0.098***	0.036***	0.040***
Panel B: R^2 from regression of outcome on portion of traditional explanatory variable related to PGSs and portion orthogonal to PGSs												
ln(Wealth)	1.72%		3.24%		0.07%		1.07%		19.95%		7.96%	
ln(Income)	1.21%		1.53%		0.04%		0.30%		7.72%		3.76%	
Education	2.82%		3.12%		0.07%		0.28%		8.95%		4.53%	
Cognition	1.76%		2.61%		0.15%		0.33%		6.92%		3.29%	
Trust	1.27%		2.60%		0.15%		0.72%		5.21%		2.47%	
Sociability	3.18%		0.97%		0.04%		0.16%		3.60%		1.67%	
Optimism	5.01%		2.39%		0.13%		0.44%		5.01%		2.29%	
Poor	0.41%		0.42%		0.05%		0.21%		2.72%		0.84%	
Height	0.40%		0.14%		0.04%		0.01%		0.30%		0.07%	
BMI	0.29%		0.05%		0.01%		0.07%		1.23%		0.35%	
Health	3.04%		1.79%		0.10%		0.55%		6.38%		2.02%	
Panel C: Fraction of regression R^2 explained by portion of traditional explanatory variable related to PGSs												
ln(Wealth)	16.95%		13.95%		60.81%		11.38%		10.80%		13.26%	
ln(Income)	19.75%		27.61%		91.98%		25.27%		24.59%		26.95%	
Education	10.80%		16.77%		85.75%		54.55%		22.69%		22.83%	
Cognition	23.93%		20.26%		42.37%		59.64%		26.78%		31.64%	
Trust	27.35%		20.04%		37.71%		32.08%		41.18%		40.36%	
Sociability	8.44%		25.41%		90.72%		63.66%		41.42%		37.99%	
Optimism	5.61%		17.71%		33.24%		41.38%		38.83%		35.38%	
Poor	78.89%		80.46%		90.16%		62.25%		65.20%		70.56%	
Height	16.28%		84.12%		5.39%		91.83%		71.23%		43.72%	
BMI	5.18%		67.09%		52.33%		29.36%		40.88%		72.10%	
Health	11.17%		20.96%		45.95%		22.84%		30.56%		44.52%	
Average	20.39%		35.85%		57.86%		44.93%		37.65%		39.94%	