Genetic Endowments, Income Dynamics, and Wealth Accumulation Over the Life-Cycle*

Daniel Barth[1], Nicholas W. Papageorge[2,3,4], Kevin Thom[5], and Mateo Velásquez-Giraldo[1]

[1]Federal Reserve Board of Governors
[2]Johns Hopkins University
[3]IZA
[4]NBER
[5]University of Wisconsin - Milwaukee

We develop and estimate a model that incorporates the influence of genetics on multiple dimensions of wealth accumulation. We study an index of genetic traits—known as a polygenic score—that strongly associates with wealth at retirement. Labor earnings and returns earned on risky assets explain more than 80% of the association between the polygenic score and wealth, with returns being particularly important. Analyses of two counterfactual policies that reduce retirement benefits show that the association between genes and wealth changes with the policy environment, but changes in this association can obscure shifts in the relationship between genes and welfare.

Keywords: Wealth, Inequality, Genetic Endowments, Saving and Portfolio Choices, Retirement.
JEL Classification: D14, D15, D31, D63, G50, H31, H55, I38, J24, J26.

---

*First draft: July 1, 2022. Current draft: October 11, 2024. We are grateful for helpful comments from Christopher Carroll, Michael Keane, and Jonathan Skinner, along with seminar participants at the University of St. Andrews, the Board of Governors at the Federal Reserve, Wisconsin - Madison, Wisconsin - Milwaukee, Banco de la República de Colombia, Yale, SITE, NBER-Aging 2022, Houston, Purdue, Miami, Rosario, Cornell, VCU, LAMES 2023, and Aarhus, along with two conferences on genes and economic analysis: Genes, Social Mobility, and Inequalities across the Life-Course and Frontiers in Economic Analysis with Genetic Data. We also acknowledge excellent research assistance from Matthew Gonzalez. Kevin Thom acknowledges funding from the National Institute on Aging of the National Institutes of Health (R01AG055654 and R56AG058726).Views and opinions expressed are those of the authors and do not necessarily represent official positions or policy of the Federal Reserve Board of Governors.

†dannybarth@gmail.com, papageorge@jhu.edu, kthom.work@gmail.com, mateo.velasquezgiraldo@frb.gov.

1 Introduction

The sources and dynamics of wealth inequality have received considerable attention in the past decade (Saez and Zucman, 2016; Catherine, Miller, and Sarin, 2020; Sabelhaus and Volz, 2022; Auten and Splinter, 2024). Recent research has established significant heterogeneity across multiple drivers of wealth accumulation, including labor market earnings, intergenerational transfers, and returns to wealth (Adermon, Lindahl, and Waldenström, 2018; Fagereng et al., 2020; Bach, Calvet, and Sodini, 2020). This heterogeneity underpins differences in wealth both across and within countries (Badarinza, Campbell, and Ramadorai, 2016).

A growing literature demonstrates that differences in individual genetic factors, and their interaction with environments, explains some of the individual variation in wealth. Traditional approaches decompose wealth variation into genetic and non-genetic components, usually by exploiting data on identical and fraternal twins (Cronqvist and Siegel, 2015) or between biological children and adoptees (Black et al., 2020; Fagereng, Mogstad, and Rønning, 2021). These methods estimate a total effect of genes on wealth, but are not well suited to identify the mechanisms that generate gene-wealth gradients. Recent advances in the collection of molecular-level genetic data and their inclusion in socioeconomic surveys have allowed for a more detailed analysis of potential mechanisms, both through documenting reduced-form associations with variables of economic importance and through the responses of these associations to changes in environments (referred to as gene-by-environment or $G \times E$ interactions). Yet, to date, genetic data have not been incorporated into models of wealth accumulation that are able to quantify the relative contribution of specific mechanisms, evaluate their potential compounding effects, and predict endogenous optimal behavior, outcomes, and welfare in counterfactual policy environments.

In this paper, we make use of molecular-level genetic variation to study the underlying mechanisms that link a particular set of genetic endowments to wealth. Our measure of genetic variation is a linear index of genetic markers, known as a *polygenic score*, associated with years of schooling, and which has previously been shown to strongly associate with wealth at retirement, labor earnings, stock market participation, and other economically relevant outcomes. We develop and estimate a structural model of consumption, savings, portfolio choices, and retirement that explicitly allows the polygenic score to affect wealth accumulation through multiple channels beyond completed education, including stock market participation costs, the return earned on stock investments, labor income, and the disutility of work. Each is motivated by reduced-form associations with the polygenic score documented in prior studies.[1] Further, we explicitly model childhood socioeconomic status (SES), later-life

[1]See Barth, Papageorge, and Thom (2020); Papageorge and Thom (2020).

inheritances, and their correlations with the polygenic score to address confounding family environments. The model also features a realistic representation of financial resources in retirement, specifically defined-benefit pensions and social security payments. We use this feature of the model to study how the association between genetic endowments and wealth changes under environments with less generous retirement benefits.

We estimate the model using moments related to income, portfolio choices, retirement decisions, and wealth. Our data come from the Health and Retirement Study, a panel data set based on surveys of elderly adults in the U.S. The estimated model parameters indicate that the association between the polygenic score and wealth is primarily driven by two channels: the returns earned on risky assets and labor income. Consistent with the recent literature, we find that heterogeneity in returns is particularly important.[2] The expected annual log-returns for those with a polygenic score one standard deviation above the mean are 2.49 percentage points higher than for individuals with a score one standard deviation below the mean. A large fraction of this unconditional relationship is not explained by educational attainment and childhood SES, which are both correlated with the EA score. The strong association between the polygenic score and expected returns induces a positive correlation between the polygenic score and stock market participation, despite model estimates that indicate stock market participation costs are unrelated to the polygenic score, similar to the findings in Calvet, Campbell, and Sodini (2007). This indirect effect on stock market participation further magnifies the association between genes and wealth driven by the returns earned on risky assets. In total, the model implies the slope of the relationship between the polygenic score and wealth would be 43% smaller if the score had no association with risky asset returns other than through its correlation with education and childhood SES.

We also find an effect due to labor market earnings, as documented by Papageorge and Thom (2020). A one standard deviation increase in the polygenic score is associated with a 6 percentage-point increase in income at age 45. Without this effect, the slope of the relationship between the polygenic score and wealth would be 25% smaller. These results highlight how the genetic endowments studied here have a compounding effect over the life-cycle—the same genetic factors that give individuals an advantage in acquiring educational attainment also increase income, and these advantages are jointly magnified by higher portfolio returns over the life-cycle.

Next, we turn to an assessment of the sensitivity of the gene-wealth association to alternative environments. This exercise relates to the burgeoning literature on how genetic associations are modified by different environments, commonly referred to as

[2]See Lusardi, Michaud, and Mitchell (2017); Bach, Calvet, and Sodini (2020); Fagereng et al. (2020); Ozkan et al. (2023).

gene-by-environment ($G \times E$) interactions. However, such analyses are nearly without exception descriptive, which limits our understanding of the importance of genes in counterfactual settings that may, but have not yet, occurred. We contribute to this literature by using our model to predict how the gene-wealth gradient would respond to environments with more or less generous social safety nets. A natural candidate for this exercise is the Social Security system in the United States, which provides significant financial resources in retirement for many households and may undergo significant modifications in the coming years. We emphasize that our structural model permits an evaluation of the gene-wealth gradient under *unobserved* counterfactual environments, such as reduced Social Security payments, which may ultimately be of particular interest to policymakers. Such an analysis is not feasible in other empirical designs, such as twin studies, which require variation in *observed* environments.

We consider two counterfactuals. The first increases the age of retirement by delaying the schedule of available Social Security benefits. The second leaves the schedule in place but reduces the level of benefits. These policies exemplify two of the main types of reforms that have been considered to address fiscal challenges of an aging population (Social Security Administration, 2022b). We find that the policies differ in the margins along which they incentivize individuals to adjust; the shift in the schedule of benefits pushes agents to extend their working lives, and the reduction in benefits incentivizes them to save more. Under the latter policy, we find that wealth inequality actually reduces between individuals with different genetic endowments. However, genetic inequality in *welfare* actually increases. This arises because individuals with lower polygenic scores earn lower returns on wealth, and under the benefit-reduction policy, these individuals must give up proportionally larger consumption to finance the lost income in retirement. Said differently, low polygenic score households, who already consume less, are forced to cut consumption relatively more to finance the lost income in retirement, flattening the gene-*wealth* gradient but steepening the gene-*welfare* gradient.

Additional counterfactual exercises reveal the important role that differences in financial sophistication play in determining the relative welfare consequences of each Social Security reform. There is a robust literature establishing a link between genetic endowments and financial sophistication (Cesarini et al., 2009, 2010; Barnea, Cronqvist, and Siegel, 2010; Cronqvist and Siegel, 2014; Black et al., 2017). Even though the median welfare effects of the two counterfactual reforms are similar under our baseline estimates, the welfare costs of the benefit-reduction plan fall substantially in alternative scenarios where individuals' financial inefficiencies are smaller. This suggests that a weakening association between genes and financial sophistication may alter the welfare consequences of potential policies.

Our analysis connects closely to the literature on household portfolio choice and heterogeneity in wealth (see Gomes, Haliassos, and Ramadorai, 2021, for a review).

Differences in financial sophistication can generate large differences in wealth across households over time (Lusardi, Michaud, and Mitchell, 2017; Campbell, Ramadorai, and Ranish, 2019). Investigating the sources of such differences and how they affect choices is crucial to understanding wealth inequality. The workhorse models in the portfolio choice literature prescribe that every household should allocate a substantial fraction of their wealth to stocks (Merton, 1969; Samuelson, 1969; Cocco, Gomes, and Maenhout, 2005). Given the generally low rates of stockholding documented in various countries (Badarinza, Campbell, and Ramadorai, 2016), later studies have incorporated financial costs to stockholding, heterogeneity in preferences, and tail-risks as ways to reconcile model predictions with household decisions (see e.g., Vissing-Jorgensen, 2002; Gomes and Michaelides, 2005; Fagereng, Gottlieb, and Guiso, 2017; Catherine, 2021). Other studies have documented heterogeneity in the returns to wealth and financial sophistication, with some showing that sophistication covaries with characteristics like education (Calvet, Campbell, and Sodini, 2007; Campbell, Ramadorai, and Ranish, 2019; Fagereng et al., 2020). We contribute to the literature by presenting evidence that the polygenic score we study can be related to financial sophistication, which can lead to compounding effects of genetic endowments over the life-cycle. Further, studies such as Fagereng et al. (2020); Ozkan et al. (2023) document intergenerational persistence in returns to wealth. Because genetic endowments are directly transfered from parents to offspring, they represent one possible explanation for this persistence.

Our paper is also related to the literature that studies genetic associations and $G \times E$ interactions using polygenic scores, surveyed by Biroli et al. (2022). A series of influential genome-wide association studies (GWAS) have estimated associations between individual genetic markers and complex socioeconomic outcomes, allowing for the construction of polygenic scores for these traits (Rietveld et al., 2013; Okbay et al., 2016; Lee et al., 2018; Karlsson Linnér et al., 2019; Hill et al., 2019). Polygenic scores for educational attainment have received particular attention, with several papers documenting associations between such scores and completed education, income, wealth, and other measures of socioeconomic success (Belsky et al., 2016, 2018; Barth, Papageorge, and Thom, 2020; Papageorge and Thom, 2020). Rustichini et al. (2023) estimate a model of intergenerational transmission that separates the influence of a polygenic score for education into channels running through cognitive and non-cognitive skills. Past studies have also found that relationships between these scores and educational attainment are moderated by environmental factors including compulsory schooling laws, school quality, birth order, and family socioeconomic status (Barcellos, Carvalho, and Turley, 2021; Arold, Hufe, and Stoeckli, 2022; Trejo and Domingue, 2018; Muslimova et al., 2020; Ronda et al., 2022; Papageorge and Thom, 2020; Belsky et al., 2016). Other studies examine interactions between environmental factors and polygenic scores for other outcomes including alcohol use, smoking, heart disease, and depressive symptoms (Fletcher and Lu, 2021; Bierut et al., 2023; Baker et al., 2022; Furuya et al., 2022). We contribute to this literature by incorporating genetic variation captured by a polygenic score into a structural life-cycle model with multiple mechanisms linking genes and wealth. This allows us to quantify these channels and perform *ex-ante* counterfactuals to assess the likely $G \times E$ effects of policies that have been proposed but not implemented. Importantly, our analysis highlights that the $G \times E$ interactions in outcomes uncovered by this literature might be different than $G \times E$ interactions in welfare and that economic theory is useful in distinguishing between these cases.

Our paper further engages the literatures on the heritability of economic outcomes and causes of intergenerational persistence in wealth. A large literature uses twin and adoption studies to demonstrate that genetic variation has explanatory power for outcomes like earnings, risk-taking and giving, investment decisions and biases, and saving decisions (Taubman, 1976; Cesarini et al., 2009, 2010; Cronqvist and Siegel, 2014, 2015; Black et al., 2020; Fagereng, Mogstad, and Rønning, 2021). Our analysis is not well suited to answer questions about the overall extent of heritability of wealth outcomes. In general, polygenic scores tend to explain only a fraction of the variance that these other methods attribute to genetic factors (Becker et al., 2021). Rather, our goal is to take advantage of the fact that polygenic scores are observed in rich longitudinal data sets like the HRS to more easily analyze the mechanisms through which they operate and how they interact with environments. Such findings may be particularly important in the literature that models intergenerational persistence in wealth and the returns to wealth (Benhabib, Bisin, and Zhu, 2011; Benhabib, Bisin, and Luo, 2019; Gayle, Golan, and Soytas, 2022; Rustichini et al., 2023; Collado, Ortuño-Ortín, and Stuhler, 2023).

The remainder of this paper is organized as follows. Section 2 describes our data and presents motivating summary statistics. Section 3 presents our model and outlines our estimation strategy. Section 4 presents parameter estimates, the fit of the empirical patterns that we target, and the implications of the estimates. Section 5 uses the model to assess how socioeconomic outcomes, lifetime welfare, and their relationship with the EA score would change under two cost-saving changes to the Social Security system. Section 6 concludes.

2 Data

2.1 The Health and Retirement Study and Our Sample

Our empirical analysis uses data from the Health and Retirement Study (henceforth HRS). The HRS surveys a representative sample of more than 20,000 Americans over the age of 50 and their spouses. The longitudinal design of the survey features biennial waves starting in 1992 and continuing until the present, which provide information on respondents' labor supply, income, wealth, financial decisions, retirement, mortality, and inheritances. Retrospective survey questions ask about childhood socioeconomic status and parental characteristics.[3] The HRS data can also be linked to Social Security Administration records to provide data on earned income throughout the lifecycle. Crucially, the HRS contains genetic information on over 18,000 respondents collected from 2006 onward. Genetic data allow us to construct various measures of genetic endowments from the behavioral-genetics literature for HRS respondents, and to study their associations with other HRS variables. These measures include polygenic scores, which are summary indices of genetic variants that have been shown to predict observable outcomes.

This paper uses a polygenic score for educational attainment developed by Lee et al. (2018) as a measure of genetic endowments that influence various dimensions of human capital.[4] We refer to it as the *EA score*. This individual-level measure aggregates genetic variants that have been linked to educational attainment. It is standardized to have a mean of 0 and a variance of 1. A relatively high EA score indicates that an individual possesses relatively more of the genetic variants that have been empirically linked out-of-sample (in our case in non-HRS samples) to educational attainment. Previous studies have shown that the EA score is predictive not only of completed education but also of other economic outcomes such as labor supply, income, and wealth, even after flexibly controlling for educational attainment (Papageorge and Thom, 2020; Barth, Papageorge, and Thom, 2020).

Earlier papers discuss a number of issues related to the empirical use of the EA score in economic analysis. Appendix A of this paper provides an overview of how polygenic scores like the EA score are constructed, important limitations in their interpretation and use, and previous research in the social sciences that has used the EA score. Some of the issues detailed there place restrictions on our sample. Accordingly, we begin with the sample used in Barth, Papageorge, and Thom (2020), which includes only households with members of European ancestry. As explained in Martin et al. (2017), a set of technical issues means that the incorporation of non-European households into our analysis would be misguided and could generate misleading conclusions about cross-ethnic group genetic differences. The sample is also limited to households with non-missing data on key measures of interest for this

[3] The HRS also collects information on a host of factors that are omitted from our analysis, including variables on health and family structure.

[4] New polygenic scores are developed and updated for different traits as more data becomes available. In the case of educational attainment, Okbay et al. (2022) construct a polygenic score based on a larger discovery sample than that of Lee et al. (2018) (3 million vs. 1.1 million individuals) and which also improves upon its predictive power. We use the polygenic score of Lee et al. (2018) because it is the latest that has been calculated and made available for HRS respondents.

study, including wealth, stock market participation, and Social Security Administration earnings records. These requirements permit a maximum sample size of 2,590 households (5,701 household-year observations) from the overall HRS sample of over 20,000 households and over 160,000 household-year observations.

We make further sample restrictions aligned to the structural model. In particular, the model is a unitary household model that abstracts from joint labor supply decisions and marriage dynamics. We focus on households that i) enter the HRS panel as married or partnered two-person male-female households; ii) remain intact (except due to the death of the female head of household); iii) are not observed earning income from jobs not covered by the Social Security Administration (SSA) data; iv) receive at least 70 percent of their SSA earnings income from the male partner; and v) have non-missing genetic data for the male partner, which we take to be the genetic endowment of the household. This set of restrictions generates a main analytic sample of 870 households with wealth data observed for a total of 2,318 household-year observations. We note, however, that to generate income and inheritance moments, we use slightly different samples. Samples used for the inheritance and income processes are described in Appendix B.1.

We use a comprehensive measure of household wealth that includes the net value of financial assets (cash, checking and saving accounts, certificates of deposits, stocks, bonds, mutual funds, trusts, and others, minus the value of non-housing debt); the net value of housing and businesses; and the balances of retirement accounts such as 401k and Keogh accounts. This is similar to the measure of wealth used in Barth, Papageorge, and Thom (2020) with the exception that we exclude the present value of Social Security payments, defined benefit pensions, and annuity income. We take the values of these components of wealth from the RAND HRS "Detailed Imputations" files. Our analysis also uses a binary indicator of whether households own any stocks, which takes a value of one for direct holdings, mutual fund holdings, and holdings through retirement accounts.

For our measure of labor income, we use the Respondent Cross-Year Summary Earnings data set of the HRS, which contains earnings data from the Social Security Administration's Master Earnings File (MEF). We use individuals' total earnings from the MEF, which include "regular wages and salaries, tips, self-employment income, and deferred compensation" (see Olsen and Hudson, 2009, for a detailed description of the components and history of measured earnings in the MEF). The earnings are derived from tax filings and are available in the Summary Earnings data set at an annual frequency, starting from the year 1951. Because the earnings data in the MEF were initially collected with the purpose of calculating Social Security benefits, earnings are top-coded at their maximum taxable level, which changes every year. As in Barth, Papageorge, and Thom (2020), we use data from the Current Population Survey to replace top-coded amounts with average earnings conditional on earning at

Table 1: Summary Statistics.

A. Demographics, Income and Wealth				B. Childhood Socioeconomic Status (SES)			
Variable	Mean	SD	N	Variable	Mean	SD	N
Birth Year	1939.5	5.81	870	Mother's Educ.	10.52	2.97	766
College	0.30	0.46	870	Father's Educ.	10.20	3.58	766
Retired				*Family SES*			
50-62	0.17	0.37	3,350	Well Off	0.07	.	766
63-67	0.59	0.49	1,786	Average	0.67	.	766
68-72	0.78	0.42	1,393	Poor	0.24	.	766
73+	0.86	0.35	503	Varied	0.01	.	766
Total Prime-Age Income (× $1,000)				*Father's Job*			
Mean	1,984.48	.	870	Manager / Prof.	0.18	.	766
Std. Dev.	802.85	.	870	Sales	0.07	.	766
25th Percentile	1,451.90	.	870	Clerical	0.03	.	766
50th Percentile	1,928.22	.	870	Service	0.04	.	766
75th Percentile	2,456.86	.	870	Manual / Operators	0.64	.	766
Household Wealth (× $1,000)				Armed Forces	<0.01	.	766
Mean	716.34	.	2,318	Don't Know	0.03	.	766
Std. Dev.	926.29	.	2,318	Missing	<0.01	.	766
10th Percentile	52.12	.	2,318	*Child Health*			
25th Percentile	150.90	.	2,318	Excellent	0.55	.	766
50th Percentile	372.03	.	2,318	Very Good	0.27	.	766
75th Percentile	872.71	.	2,318	Good	0.13	.	766
90th Percentile	2,871.72	.	2,318	Fair	0.03	.	766
Any Stocks	0.68	0.47	2,318	Poor	0.01	.	766

Summary statistics for demographics, income, wealth, and SES variables in our main analytical sample.

least the maximum taxable amount for a given year.

Table 1 presents basic descriptive statistics for the analytic sample. Approximately 30 percent of the sample has a college degree, making the sample more highly educated than the overall HRS population. Retirement rates increase from approximately 17 percent for ages 50-62 to about 59 percent for ages 63-67. By age 73, 86 percent of the sample has retired. Table 1 also presents basic descriptive statistics on the log of total prime-age SSA earnings. To construct this variable, we sum male earnings from age 30 to 60. The median of total income over this age range is $1.9 million, which would constitute an average of approximately $59,000 per year in 2010 dollars. Table 1 also provides detailed descriptive statistics for wealth for household-year observations in which the male household member was aged 60-70. The median wealth in the sample is approximately $372,000.

2.2 Genetic Endowments and Family Environments

A natural concern is that genetic endowments are endogenous to family environments. Parents who provide their children with genetic material also provide them with family environments, including resources that could benefit their educational attainment, labor market behaviors and outcomes, financial decision-making, and wealth accumulation. As many such factors are likely to be unobserved and thus omitted from empirical analyses, estimated coefficients relating the EA score to outcomes, including education, are likely to be upwardly biased. A host of studies have employed different methods to address this concern. For example, Trejo and Domingue (2018) and Belsky et al. (2018) rely on within-sibling variation in the EA score. This amounts to adjusting for a family fixed effect, and nearly all relationships hold despite some differences. Another method is to control for a rich set of variables that describe childhood environments. Both Ronda et al. (2022) and Arold, Hufe, and Stoeckli (2022) show that controlling for observed measures of family background can reduce the bias substantially. Since the HRS does not have the data to perform a within-family analysis, our approach is to incorporate information on childhood SES to control for key dimensions of childhood socioeconomic status. In what follows, we discuss how we summarize this information into a single variable.[5]

To construct a summary SES measure for the male members of each household in our sample, we estimate a cross-sectional regression of the following form:

$$\mathrm{Educ}_i = b_0 + b_1 \mathrm{EA}_i + b_2 X_i + e_i \tag{1}$$

Here X_i contains a vector of background variables that includes: mother's years of schooling, father's years of schooling, dummy variables for different subjective assessments of family SES growing up, dummy variables for categories of father's occupation growing up, and dummy variables for subjective assessments of health in childhood. Table 1 presents descriptive statistics on the components of X_i. After estimating the above equation, we construct an index of childhood SES as: $\text{SES Score}_i = \widehat{b_2} X_i$. We normalize this measure so that it has a mean of zero and a unit standard deviation. While we do not have exogenous variation in the polygenic score EA_i in our sample, the results from Ronda et al. (2022) and Arold, Hufe, and Stoeckli (2022) give us some reason to believe that once we condition on SES Score_i, the associations we observe between EA_i and human capital outcomes may be close to causal effects.

[5] In section 3, we specify the structural model and discuss how this variable enters through unobserved heterogeneity that is correlated with the EA score.

2.3 Descriptive Associations

Table 2 presents basic regressions that highlight strong empirical relationships between the EA Score, wealth, and stock market participation, which help to motivate our structural model. Panel A presents regressions of log household wealth on explanatory variables, including the EA Score for household-year observations where the male household member is aged 60-70.[6] Column (1) includes the EA Score, the SES Score, and a dummy variable for college education as controls. The coefficients on all three variables are substantial and statistically significant. The results suggest that a one standard deviation higher EA Score is associated with 27 percent higher wealth. Column (2) adds the log of total prime-age income, and Column (3) adds a dummy variable for holding any stocks in a given household-year. Controlling for income reduces the coefficient on the EA Score to 0.20, while additionally controlling for stocks reduces the coefficient to 0.15. However, this coefficient remains significant, suggesting that in this sample, even controlling for college, childhood SES, life-time earnings, and stocks, a one standard deviation higher EA score is associated with approximately 15 percent higher household wealth. Panel B of Table 2 presents regressions of a dummy variable for holding any stocks on the EA score, college, and childhood SES. The estimates in Column (4) suggest that a one standard deviation higher EA score is associated with a 6.8 percentage point increase in the likelihood of owning stocks. This coefficient is attenuated (0.049 v.s. 0.068) but remains highly statistically significant after controlling for lifetime earnings in Column (5).

The descriptive associations presented in Table 2 are consistent with earnings and portfolio choice playing major roles in mediating the relationship between the EA Score and household wealth. The model we develop and estimate below attempts to explain these associations as arising from the effects of the EA Score on the earnings process, the fixed costs of stock market participation, returns on stock market investments, and the disutilty of labor, while accounting for family background and the inheritance process. We focus on these channels in part because of empirical results found in Barth, Papageorge, and Thom (2020). However, several other channels could theoretically link the endowments measured by the EA Score and household wealth. In Appendix B, we explore five such mechanisms: business ownership, risk preferences, fertility, marital history, and longevity expectations. Measures of all of these channels significantly predict household wealth but only modestly attenuate the EA Score's associations with wealth and stock ownership. We thus abstract from these channels in developing our structural model.

[6] To reduce omitted variables bias, it is common practice to control for the first 10 principal components of the full set of genetic variables, which helps to control for broad patterns in the genetic data that might arise from ethnic or regional differences and are omitted from the analysis. Appendix A discusses this practice in more detail.

Table 2: Summary Statistics: Mechanisms.

	Panel A: log Wealth			Panel B: Any Stocks	
	[1]	[2]	[3]	[4]	[5]
EA Score	0.276	0.203	0.152	0.068	0.049
	(0.054)	(0.050)	(0.045)	(0.015)	(0.014)
SES Score	0.294	0.222	0.169	0.069	0.050
	(0.054)	(0.058)	(0.052)	(0.016)	(0.015)
College	0.527	0.331	0.260	0.119	0.067)
	(0.109)	(0.105)	(0.095)	(0.032)	(0.030)
log Prime Inc.		0.969	0.699		0.255
		(0.195)	(0.180)		(0.030)
Any Stocks			1.058		
			(0.093)		
N	2259	2259	2259	2259	2259

This table reports results from models predicting log Wealth (Panel A) and a dummy variable for any stocks (Panel B). All regressions include the following controls: the first 10 principal components of the genetic data, an indicator for missing SES scores, and interactions between this missing SES indicator and any featured explanatory variables (the EA Score, College, log Prime Income, or Any Stocks if they are present in the specification).

3 Model

The model features heterogeneous agents that live from age 21 to a maximum age of 90. Each year, agents decide how much wealth to consume, how to allocate savings between a risky and a risk-free asset, and—once old enough—whether to retire or continue working for another year. Below we describe each component of the agents' dynamic problem.

3.1 Ex-Ante Heterogeneity

Agents are indexed by subscript i. They enter the model with four dimensions of observable heterogeneity: birth year (BY_i), an indicator for college completion (Coll_i), the EA score (EA_i), and an indicator equal to one if the agent participates in a defined benefit pension plan (DB_i). We allow these characteristics to influence agents' income, utility cost of work, expected returns on the risky asset (stocks), and the cost of stock market participation.

We also allow unobserved heterogeneity to affect agents' earnings, stock market participation costs, and expected stock returns. We model these three dimensions of heterogeneity as a vector of individual-specific "fixed effects" $\vec{\zeta_i} = [\zeta_i^w, \zeta_i^F, \zeta_i^R]'$. These parameters account for heterogeneity along dimensions that we do not model directly.

One particularly important source of heterogeneity is childhood socioeconomic status and rearing environment. If the EA score in part represents genetic endowments that are conducive to achieving higher levels of income and financial sophistication, then we would expect the average-, low-, and high-EA individuals to grow up in different environments. This indirect channel, known as genetic nurture, could lead us to overestimate the influence of the EA score (Kong et al., 2018; Young et al., 2018; Ronda et al., 2022). To address this issue, we model unobserved heterogeneity as a combination of a fully random component ($\vec{\mathcal{Z}}_i$) and a component that is correlated with the EA score through childhood socioeconomic status (SES_i),

$$\underbrace{[\zeta_i^w, \zeta_i^F, \zeta_i^R]'}_{\vec{\zeta}_i} = \underbrace{[z_w, z_F, z_R]'}_{\vec{z}} \times \mathrm{SES}_i + \underbrace{[\mathcal{Z}_i^w, \mathcal{Z}_i^F, \mathcal{Z}_i^R]'}_{\vec{\mathcal{Z}}}, \qquad \vec{\mathcal{Z}} \sim \mathcal{N}\left(\vec{0}, \Sigma_{\mathcal{Z}}\right) \tag{2}$$

where $\Sigma_{\mathcal{Z}} = \mathrm{diag}[\sigma^2(\mathcal{Z}^w), \sigma^2(\mathcal{Z}^F), \sigma^2(\mathcal{Z}^R)]$, and $\mathrm{SES}_i = \phi EA_i + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Here $\vec{z}$, $\Sigma_{\mathcal{Z}}$, ϕ and σ_ε^2 are parameters to be estimated.

3.2 Utility

A surviving agent i in period t derives utility from consumption and leisure through the utility function

$$u_{i,t}(C_t, \ell_t) = \frac{\left[C_t^\gamma \left(1 - 0.34 \times \ell_t\right)^{1-\gamma}\right]^{1-\omega}}{1-\omega} - d_{i,t} \times \ell_t, \tag{3}$$

where C_t is consumption and ℓ_t is a binary variable that indicates whether the agent is working ($\ell_t = 1$) or not ($\ell_t = 0$). This assumes that labor is indivisible and that it consumes 34% of an agent's endowment of time (8 of 24 hours per day), which we normalize to 1. A Cobb-Douglas function aggregates consumption and leisure with a weight $\gamma \in [0, 1]$ on consumption. The aggregate passes through a constant relative-risk-aversion function with coefficient of relative risk aversion ω. The term $d_{i,t} \times \ell_t$ is an additive utility cost of work. We allow this cost to vary with agents' education, EA score, and age:

$$d_{i,t} = d_0 + d_{\mathrm{Coll}} \times \mathrm{Coll}_i + d_{\mathrm{EA}} \times \mathrm{EA}_i + d_{\mathrm{Age}} \times \max\{\mathrm{Age}_{i,t} - 50, 0\}. \tag{4}$$

Our specification of the utility function allows for two different sources of heterogeneity in the disutility of work that could drive the empirical relationships between observable characteristics—the EA score in particular—and retirement patterns. The Cobb-Douglas aggregator in Equation 3 makes the value of leisure change with agents' consumption and, therefore, agents with different income levels (that support different consumption levels) will value leisure differently. This channel, whose importance is highlighted by Heckman (1974), is common in recent models of life-cycle labor supply (see, e.g., Low, Meghir, and Pistaferri, 2010; Blundell et al., 2016). The additive term in Equation 4 is intended to capture costs of work that are unrelated to the level consumption. The heterogeneity of this cost is intended to capture potential differences in, e.g., the physical intensity of jobs that may co-vary with observable traits. Studies like Attanasio, Low, and Sánchez-Marcos (2008); Bagchi (2015) have used specifications with both additive and multiplicative costs of labor similar to ours.

Finally, to accommodate the fact that high-income households have higher saving rates and are slow to run down their wealth at old ages (Dynan, Skinner, and Zeldes, 2004), we include a "joy of giving" bequest motive (Carroll, 2002) that produces utility from end-of-life wealth. We use the functional form of De Nardi, French, and Jones (2010), in which a person who dies with savings $S_{i,t}$ receives utility

$$\varphi(S_{i,t}) = \theta \frac{(S_{i,t} + \kappa)^{1-\omega}}{1-\omega}, \tag{5}$$

where ω is the same coefficient of relative-risk-aversion as above and θ and κ are parameters that we estimate and which govern the intensity of the motive and the degree to which bequests are a luxury good.[7]

3.3 Financial Assets

At the end of each period, agents decide how to allocate their savings between two financial assets: a risk-free bond with return factor R and a risky asset representing the stock market with an agent- and time-specific return factor $\tilde{R}_{i,t}$. Short-selling of either asset is not permitted.

For agent i, log-returns in the risky asset are

$$\ln \tilde{R}_{i,t} = \ln R_t^{\text{SP500}} - \mu^{\text{SP500}} \times g(r_0 + r_{\text{Coll}} \times \text{Coll}_i + r_{\text{EA}} \times \text{EA}_i + \zeta_i^R), \tag{6}$$

where $\ln R_t^{\text{SP500}}$ and μ^{SP500} are the log-return of the S&P 500 stock market index in year t and its mean log-return, respectively. All agents assume market returns are log-normal, $\ln R_t^{\text{SP500}} \sim \mathcal{N}(\mu^{\text{SP500}}, \sigma^{\text{SP500}})$. The factor $g(\cdot)$ captures agents' degree of inefficiency when investing in risky assets. It takes the form of a logistic function: $g(x) = \frac{e^x}{1+e^x}$, which ranges from 0 to 1. An efficient agent ($g \approx 0$) will replicate the market's log-returns. An inefficient agent ($g \approx 1$) will have expected log-returns close to 0. This parsimonious specification is similar to that of Lusardi, Michaud,

[7]A greater κ reduces the maximum marginal utility that can be gained from bequests. This increases the level of wealth at which an agent would start to sacrifice consumption for bequests (see Carroll, 2002).

and Mitchell (2017). Note that all agents face the same volatility on risky asset log-returns; only the mean log-return is agent-specific.

We conceive of heterogeneity in risky asset log-returns as proxying for sound investment decisions. Paying higher fees on mutual fund investments or excessive trading, for instance, would degrade the average return earned in the market. Inopportune market-timing strategies, such as buying during periods of high price-earnings ratios and selling during periods of low price-earnings ratios, would also be detrimental to returns. Further, lower average log-returns can result from more volatile simple returns (see Campbell, Ramadorai, and Ranish, 2019). Therefore, our model can also be interpreted as capturing differences in the effectiveness with which households diversify financial risks.

To own the risky asset, agents must pay a per-period monetary cost that represents the administrative and opportunity costs of managing investments (Vissing-Jorgensen, 2002). The cost F_i depends on *ex-ante* demographic characteristics:

$$\ln \mathrm{F}_i = f_0 + f_{\mathrm{Coll}} \times \mathrm{Coll}_i + f_{\mathrm{EA}} \times \mathrm{EA}_i + \zeta_i^F. \tag{7}$$

Capital gains are taxed at a constant rate τ^c.

3.4 Labor Income

We model pre-tax labor income $\tilde{W}_{i,t}$ as the log-sum of a deterministic component that depends on individual characteristics and aggregate trends, the fixed unobserved heterogeneity draw $\mathcal{Z}_i^w$, and an agent- and time-specific shock $\epsilon_{i,t}^w$:

$$\ln \tilde{W}_{i,t} = f(\mathrm{Age}_{i,t}, \mathrm{EA}_i, \mathrm{Coll}_i, \mathrm{SES}_i, \mathrm{DB}_i, \mathrm{Year}_t, \mathrm{Unemp}_t) + \mathcal{Z}_i^w + \epsilon_{i,t}^w, \tag{8}$$

where Unemp_t is the aggregate unemployment rate in period t. The wage shock is independent across time and agents and is normally distributed, $\epsilon_{i,t}^w \sim \mathcal{N}(0, \sigma^2(\epsilon^w))$. We present our specification and estimates for pre-tax labor income in Appendix B.1.

We use two types of labor income taxes: a constant-rate tax (as in Chai et al., 2011), and an additional proportional tax that applies up to a year-dependent maximum $\bar{T}(\mathrm{Year})$ and represents Social Security taxes.[8] Given tax rates τ^W and τ^{FICA}, we model this tax scheme with a function $\tau_t(\cdot)$ that computes post-tax income as $\tau_t(\tilde{W}) = \tilde{W}(1 - \tau^W) - \tau^{\mathrm{FICA}} \times \min\{\tilde{W}, \bar{T}(\mathrm{Year}_t)\}$.

3.5 Retirement and Social Security

Once agents reach a minimum age, they can decide to retire. This decision is irreversible and happens after receiving labor income so that it takes effect in the

[8] We use the Social Security Administration's historical maximum taxable incomes as $\bar{T}(\mathrm{Year}_t)$.

following year. We impose the restriction that the first year of retirement has to occur in an age interval $[\text{Age}_0^R, \text{Age}_f^R]$, which we set to $[62, 80]$. Agents with defined benefit pension plans start receiving their payments in the first year in which they do not work. We assume that agents start claiming Social Security benefits the moment they stop working or at the minimum claiming age ($\text{Age}_{\text{min}}^{SS}$) if they stop working before that.[9]

Consistent with the payout policy in the U.S., social security benefits are a concave function of agents' average income over their highest-earning years and increase with each additional year of work up to a maximum. Our methodology for computing benefits follows that of the Social Security Administration; we provide a detailed description in Appendix C. Our main simplification is that we use expected (rather than realized) earnings when computing SS benefits. That is, in each period t in which the agent could choose to retire, instead of the agent looking backwards at her *realized* earnings (the top 35 years of which determine SS benefits), the agent anticipates receiving SS benefits that are determined by her *expected* top-35 yearly earnings as of period t, based on the income function described in Equation (8). This allows us to avoid the additional state variable of realized cumulative earnings. Additionally, because the model has no permanent income shocks, transitory shocks over the life-cycle should largely cancel out, and expected earnings should be a good representative of realized earnings. We use $\text{SSB}_i(n)$ to denote the yearly benefits that person i would receive if she retired at age n.

In addition to social security benefits, retirees who have a defined benefit pension plan ($\text{DB}_i = 1$) receive pension income. We model the annual amount of DB pension payments (DBf_i) as a log-linear function of the time-invariant components of income, EA_i, and Coll_i, which we estimate from the data. This is similar to our treatment of social security income: to avoid the introduction of an additional state variable, we tie DB retirement benefits to the predictable component of earnings rather than realized earnings. We present the full specification and estimates of the defined benefit pension income process in Appendix B.1. Both social security benefits and defined benefit pension flows are taxed at a constant rate τ^s.

3.6 Inheritances

Every period, agents receive inheritances ($\text{Inher}_{i,t}$) that follow an agent- and age-specific stochastic process. With probability $P_i^I(\text{Age}_{i,t})$, individuals receive an inheritance in period t of amount $\text{C.Inher}_i(\text{Age}_{i,t})$. With probability $(1 - P_i^I(\text{Age}_{i,t}))$, the

[9]In our baseline scenario, the minimum claiming age and minimum retirement age are both 62, and therefore agents will always start claiming Social Security benefits the same year they stop working. This changes in our counterfactual policy experiments, one of which shifts the minimum claiming age but leaves open the possibility of agents retiring before that point.

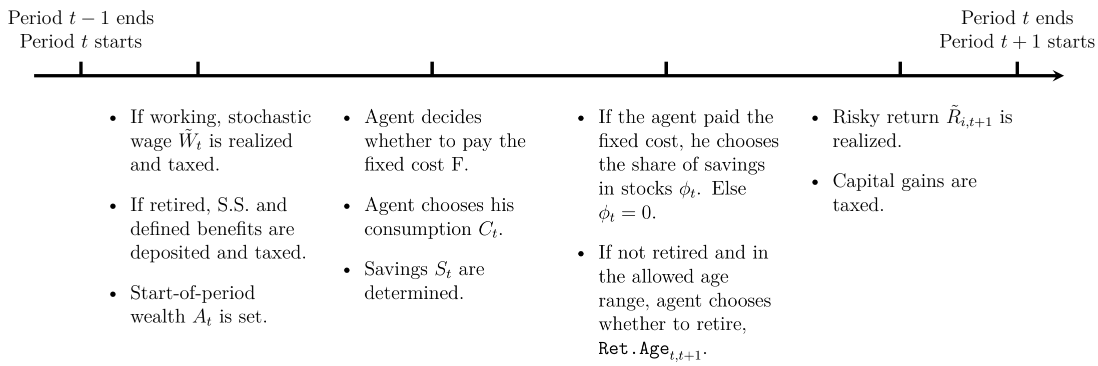


Figure 1: Timing of decisions and shock realizations.

individual does not receive an inheritance in period t. Both the probability of receiving an inheritance P_i^I and the value of inheritances conditional on reception C.Inher_i depend on Coll_i and EA_i. We estimate both functions using our HRS sample; their specification and parameter estimates can be found in Appendix B.1.

3.7 Recursive Representation and Timing Summary

Agents maximize their expected discounted lifetime utility using a discount factor β and taking into account their probability of survival δ_t. An agent's possible choice variables at a given time are his consumption $C_{i,t}$, the fraction of his savings allocated to the risky asset $\phi_{i,t}$, and his retirement status for next period. His state vector consists of his beginning-of-period wealth $A_{i,t}$ and his current retirement status. Retirement status is captured using the age at which the agent retired, $\texttt{Ret.Age}_{i,t}$. This discrete variable takes the value n if the agent retired at age n, and takes the null value $\emptyset$ if the individual has not yet retired. We track the age of retirement because it influences the level of social security benefits that agents receive. Figure 1 summarizes the timing of decisions and shocks in our model.

The choices and constraints that an agent faces depend on his age and retirement status. We illustrate the transitions of the model depicting the problem of an agent who has not retired yet ($\texttt{Ret.Age}_{i,t} = \emptyset$) but who has the option to retire. For a level of assets $A_{i,t}$, his value function is

$$
\begin{aligned}
V_{i,t}(A_{i,t}, \emptyset) = &\max_{C_{i,t}, \phi_{i,t}, \texttt{Ret.Age}_{i,t+1}} u(C_{i,t}, 1) + \beta \delta_t \mathbb{E}_t \left[V_{t+1}(A_{i,t+1}, \texttt{Ret.Age}_{i,t+1}) \right] + \not{\delta}_t \varphi(S_{i,t}) \\
0 \leq &C_{i,t}, \quad 0 \leq S_{i,t}, \quad 0 \leq \phi_{i,t} \leq 1, \quad \texttt{Ret.Age}_{i,t+1} \in \{\emptyset, \text{Age}_{i,t} + 1\}, \\
S_{i,t} = &A_{i,t} - C_{i,t} - \text{F} \times \mathbf{1}\left[\phi_{i,t} > 0\right], \\
A_{i,t+1} = &\left\{ (1 - \tau^c) \left[\phi_{i,t} \tilde{R}_{t+1} + (1 - \phi_{i,t}) R \right] + \tau^c \right\} \times S_t + \text{Income}_{i,t+1}(\texttt{Ret.Age}_{i,t+1}),
\end{aligned}
$$

where $S_{i,t}$ denotes the agent's savings and $\not{\delta}_t \equiv 1 - \delta_t$ is the probability of death. We have aggregated all sources of non-capital income for conciseness. We present disaggregated representations of all the optimization problems that agents solve at different points of their lives in Appendix D.

3.8 Estimation

Our approach to estimating model parameters relies on standard methods, so we relegate most details to Appendix B and provide a brief overview here. We estimate the model in two steps. In the first step, we directly estimate the specifications of wages (Equation 8), inheritances, and defined benefit pension flows on the HRS sample data. The details and results of the first step are found in Appendix B.1.

In the second step, we use the method of simulated moments (MSM) to estimate the parameters that govern the financial costs to stock market participation (Equation 7), inefficiencies in risky investments (Equation 6), the disutility of work (Equation 4), the dispersion of unobserved heterogeneity (Equation 2), the bequest motive (Equation 5), and the influence of childhood socioeconomic status on the unobserved heterogeneity draws for costs and returns (Equation 2).[10] The full set of parameters that we estimate internally is

$$\Theta = \{f_0, f_{\text{Coll}}, f_{\text{EA}}, \sigma(\mathcal{Z}^F), r_0, r_{\text{Coll}}, r_{\text{EA}}, \sigma(\mathcal{Z}^r), d_0, d_{\text{Coll}}, d_{\text{EA}}, d_{\text{Age}}, \theta, \kappa, z_F, z_R\}. \quad (9)$$

The algorithm proceeds as follows. A candidate set of parameters is chosen. Given these parameters, we solve the model, which delivers policy functions (mappings from state variables to choices) and transition rules (mappings from current-period state and choice variables to one-period-ahead state variables). Next, we simulate populations that match the HRS sample on observables. Given the observable dimensions along which agents can differ in our model, there are 400 potential types of *ex-ante* different agents.[11] However, only 190 of these possible combinations are actually observed in the HRS sample. For each person in our sample, we simulate 10 agents with matching characteristics. We simulate 27 such populations, one for each of the 27 potential combinations of draws of unobservable heterogeneity ($\vec{\zeta}_i$). Thus, our simulated populations have $5,130 = 190 \times 27$ types of *ex-ante* different agents. Armed with policy functions and transition rules, we can then simulate sequences of choices and outcomes, which delivers a data set that can then be compared to the actual HRS analytic sample. To compare the simulated data set to the HRS data set, we

[10] Since income is observable, we estimate the parameters pertaining to unobserved heterogeneity of income (z_W, $\sigma(\mathcal{Z}^W)$) directly. See Appendix B.1.

[11] The observable dimensions of heterogeneity are birth year (BY_i), college completion (Coll_i), EA score (EA_i), and whether the agent participates in a defined benefit pension plan (DB_i). Appendix B discusses how we form the possible groups.

compare some moments directly (e.g., the simulated and observed retirement wealth distribution). We also use indirect inference to identify parameters that are not directly observed. For example, we do not directly observe the disutility of labor. To identify this parameter, however, we can regress labor supply on observable variables in both the simulated and the HRS data set and then compare coefficients from these regressions. The moments that we directly target with our estimation routine are the mean and percentiles 10, 25, 50, 75, and 90 of the distribution of wealth between ages 60 and 70; the stock-market participation rate; and the coefficients of auxiliary regressions of log-wealth, stock ownership, and retirement on the EA score, education, our index of socioeconomic status, and other controls.

We repeat this process for many different candidate parameter sets and search for the parameter vector that minimizes the distance between the simulated and empirical moments. See Appendix B for a detailed description of the estimation procedure, non-estimated parameters, technical details, and a discussion of identification. Appendix E discusses the numerical solution of the model.

4 Results

4.1 Model Fit, Parameter Estimates, and Their Implications

Table 4 shows that the estimated model matches most of the empirical patterns remarkably well. We closely match the distribution of wealth up to the 75th percentile but under-predict wealth in the upper tail of the distribution. We also match relationships between the EA score, college attendance, wealth, and stock ownership as captured by the auxiliary regressions. The overall stock ownership rate is also close to the sample rate. For retirement, while the model fits the auxiliary regression that relates retirement to the EA score and college attendance, the age-binned retirement rates show slight discrepancies arising mainly from over-estimating the fraction of agents that retire between ages 63 and 67.

Table 3 presents parameter estimates and their standard errors. To compute standard errors, we calculate the targeted moments for 50 different bootstrapped sub-samples of our analytical sample data. Within each subsample, we select the parameters (out of a pre-specified grid of 20,000 points) that minimize the distance between the model-implied and bootstrapped moments. This gives 50 sets of parameter vectors, one for each subsample, each of which minimizes the loss function within that particular subsample. We compute standard errors as the standard deviation of each parameter over the 50 resulting parameter vectors.[12]

[12] The 20,000 parameter vectors are the initial grid that we use in our main estimation routine.

Table 3: Internally-estimated parameters.

Participation cost $\ln \mathrm{F}_i = f_0 + f_{\mathrm{Coll}} \times \mathrm{Coll}_i + f_{\mathrm{EA}} \times \mathrm{EA}_i + \zeta_i^F$			**Risky asset returns** $\mathrm{Inefficiency}_i = g(r_0 + r_{\mathrm{Coll}} \times \mathrm{Coll}_i + r_{\mathrm{EA}} \times \mathrm{EA}_i + \zeta_i^R)$		
f_0	f_{Coll}	f_{EA}	r_0	r_{Coll}	r_{EA}
-0.9867	0.0311	0.0066	-0.0366	-1.1055	-0.6610
(0.2092)	(0.0369)	(0.0143)	(0.0608)	(0.2568)	(0.1326)

Disutility from work $d_{i,t} = d_0 + d_{\mathrm{Coll}} \times \mathrm{Coll}_i + d_{\mathrm{EA}} \times \mathrm{EA}_i + d_{\mathrm{Age}} \times \max\{\mathrm{Age}_{i,t} - 50, 0\}$			
d_0	d_{Coll}	d_{EA}	d_{Age}
0.3961	-0.0052	-0.0033	-0.0241
(0.0816)	(0.0015)	(0.0009)	(0.0062)

Unobserved heterogeneity $\vec{\zeta}_i = \vec{z} \times \mathrm{SES}_i + \vec{\mathcal{Z}}_i$				**Bequest motive** $\varphi(S_{i,t}) = \theta(S_{i,t} + \kappa)^{1-\omega}/(1-\omega)$	
$\ln \sigma(\mathcal{Z}^F)$	$\ln \sigma(\mathcal{Z}^r)$	z_F	z_R	$\ln \kappa$	$\ln \theta$
1.1838	-4.0326	-0.0434	-0.7250	7.0638	6.9423
(0.5698)	(1.3845)	(0.0475)	(0.1451)	(0.2474)	(0.5353)

This table presents parameter estimates from the method of simulated moments. See the main text for details about the model and targeted moments. Standard errors are reported in parentheses and calculated using a bootstrap approximation that is also discussed in the main text.

Table 4: Targeted moments in the HRS and in the estimated model.

Wealth distribution
Thousands of dollars

	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	Mean
Data	52	151	372	873	1819	716
Model	52	166	427	822	1379	605

Wealth regression
Coefficients from: $\ln \mathrm{Wealth}_{i,t} = X_{i,t}\beta + \varepsilon_{i,t}$

	EA	SES	Coll	Stocks
Data	0.15	0.17	0.26	1.06
Model	0.17	0.18	0.26	1.12

Stock-ownership regression
Coefficients from: $\mathrm{Stocks}_{i,t} = X_{i,t}\beta + \varepsilon_{i,t}$

	EA	SES	Coll	ln E. Inc.
Data	0.049	0.050	0.067	0.255
Model	0.046	0.052	0.069	0.220

Retirement regression
Coefficients from: $\mathrm{Retired}_{i,t} = X_{i,t}\beta + \varepsilon_{i,t}$

	EA	Coll
Data	-0.0059	-0.0564
Model	-0.0059	-0.0575

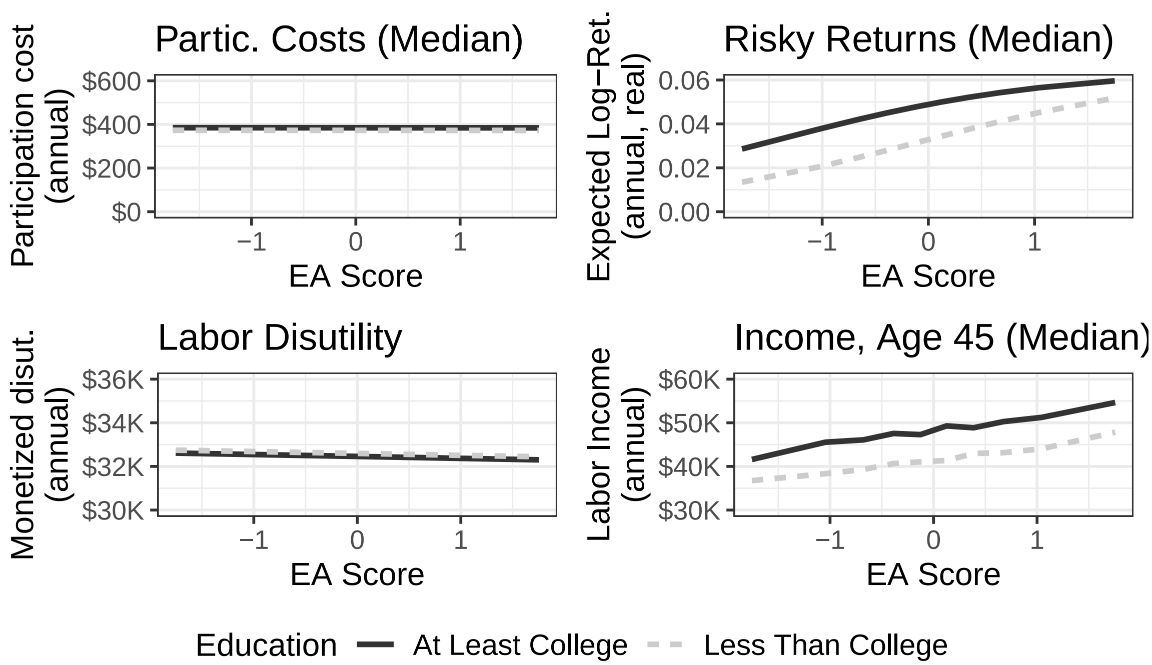
Retirement rates
Fraction of individuals retired by age bracket

	$[50, 62]$	$[63, 67]$	$[68, 72]$	≥ 73
Data	0.17	0.59	0.78	0.86
Model	0.11	0.69	0.91	1.00

Stock ownership rate
Ages 60 to 70

	Rate
Data	0.68
Model	0.72

This table reports the moments that we target in estimation, calculated both using our HRS analytical sample and using simulations from the model with the estimated parameter values. See the main text for variable definitions, sample descriptions, and complete specifications of the controls included in $X_{i,t}$ for each regression.



The disutility of labor is taken at age 50 and monetized (see the main text for details). We find the participation costs, expected log returns, and disutility from work at age 50 for each of our 5,130 agent types and weight them by the number of agents of each type in our simulated population. The figure presents medians at different levels of the EA score and educational attainment. Labor disutility depends only on agent's age, EA score, and educational attainment, so there is no remaining heterogeneity after conditioning on these characteristics. For income, we report the median of EA score-education combinations at age 45 from our simulations.

Figure 2: Economic fundamentals, the EA score, and education.

According to estimates in Table 3, both the EA score and education have negligible effects on the cost of stock market participation, F_i. This is shown in Figure 2, which plots the median participation costs, expected risky returns, disutility of labor, and income at age 45 against the EA score for both levels of education. The top-left panel shows no meaningful variation of the median annual participation cost with either the EA score or education. We estimate this cost to range between $372 and $385, which is similar to estimates in recent studies in household finance (see e.g., Fagereng, Gottlieb, and Guiso, 2017; Catherine, 2021).

Risky asset returns, however, appear to be significantly affected by both the EA score and education. Table 3 and the top-right panel of Figure 2 show that, between the lowest and highest EA deciles, the median expected log-return on stocks increases from 1.3% to 5.2% for those without a college degree and from 2.8% to 6.0% for those with a college degree. The direction of these estimated relationships is consistent with past studies (e.g., Calvet, Campbell, and Sodini, 2007; Fagereng et al., 2020), which find that returns to wealth are heterogeneous and correlated with wealth, income, and education. Each of these factors are correlated with the EA score, suggesting the potential for a genetic mechanism underlying those relationships. The specification of log-returns permits two potential sources of heterogeneity that would produce higher mean log returns: a higher mean gross return or lower return variance. This is because the mean log return is approximated by the mean gross return minus one-half the gross return variance. Our model does not allow us to separately identify these sources, however, Campbell, Ramadorai, and Ranish (2019) find that most of the difference in the mean log return across households is due to heterogeneity in the variance of simple returns rather than heterogeneity in the mean simple return.

Part of the relationship between the EA score and returns that is depicted in Figure 2 is due to the relationship between the EA score and SES.[13] To evaluate the size of the estimated association that is not due to SES—the part that operates through r_{EA} in Equation 6—we calculate expected risky log-returns of our simulated population of agents assuming their EA scores were one standard deviation (1.0) higher, holding their other characteristics, including SES, constant. This is essentially the partial derivative of returns with respect to the EA score, computed separately for each agent. Over the full simulated population, this produces an average increase in expected log returns of 0.78 percentage points. That is, on average, holding all else equal, a one standard deviation higher EA score results in a 0.78 percentage points higher annual log return.

The estimated degree of heterogeneity in returns is plausible and broadly consistent with the findings of studies that have measured the distribution of returns and Sharpe ratios in other countries, or modeled financial proficiency as an endogenous investment. By design, the functional form of the expected log-return of the risky asset (Equation 6) forces it to be between 0 and the benchmark μ^{SP500} for every agent. A similar range is used in studies such as Lusardi, Michaud, and Mitchell (2017).[14] Our estimates distribute agents across this predetermined range. Appendix F presents detailed statistics of the distribution of Sharpe ratios implied by our estimates. Households in our estimated model attain Sharpe ratios that can range from 36% of that of the market benchmark in the 25th percentile to 76% in the 75th percentile. These ranges are plausible when compared with those measured by, e.g., Calvet, Campbell, and Sodini (2007) in Sweden and Gaudecker (2015) in the Netherlands. They are also consistent with the implications of the model developed by Lusardi, Michaud, and Mitchell (2017) in which, at age 50, the average agent earns between 43% and 67% of the market's log-return premium, depending on his education.

[13]Individuals with higher EA scores have a higher expected SES (see Section 3.1). This shifts their expected draws of unobserved heterogeneity (see Equation 2) and, in turn, their expected returns.

[14]A minor difference is that the lower limit for expected log-returns in Lusardi, Michaud, and Mitchell (2017) is set to the log risk-free rate instead of 0.

Similar to previous studies, our income estimates imply a positive and significant association between the EA score and labor income even after controlling for educational attainment. The bottom-right panel of Figure 2 depicts the median labor income of agents with different EA scores and levels of education at age 45, which is close to their peak for those without a college degree. At this age, the median earnings for those without a college degree in the 4th and 7th decile of the EA score are \$40,700 and \$43,000, respectively. For those with a college degree the median earnings at this age and the same EA score deciles \$47,500 and \$48,900 respectively. Our estimates of the income process are reported in Appendix B.1.

The estimated additive component of disutility from labor varies little with education, the EA score, and age. We monetize the disutility to convey its magnitude. For a utility cost $d_{i,t} > 0$, we find the monetary value $m_{i,t}$ that an agent with no dynamic considerations and a baseline consumption of \$40,000 would be willing to pay to avoid the cost $d_{i,t}$.[15] Formally, $m_{i,t}$ solves

$$u_{i,t}(C_t = \$40k, \ell_t = 1) = u_{i,t}(C_t = \$40k - \$m_{i,t}, \ell_t = 1) + d_{i,t}.$$

The bottom left panel of Figure 2 displays monetized costs of work $m_{i,t}$ at age 50 against the EA score for both education groups. The figure confirms that there is no meaningful heterogeneity in our estimated disutility of work.

The disutility of work parameter would most clearly affect retirement. If wealth and expected retirement income are held constant, a greater disutility of work implies earlier retirement rates. The null results for the disutility of work suggest the model can fit the observed retirement patterns based on variation in wealth, income, and expected retirement income, without needing to assign additional power to a distaste for work. This finding is also supported by the strong fit of the retirement regression moments.

4.2 Life-Cycle Choices and Outcomes

The estimates suggest no meaningful relationship between the EA polygenic score and either stock market participation costs or disutility from work, but do suggest sizable associations with labor incomes and risky-asset returns, even at comparable levels of education. This section examines how these differences influence other choices and outcomes such as consumption, retirement, stock-market participation, and wealth across the life cycle. Figure 3 presents the simulated age-profiles of these choices and outcomes for agents in different deciles of the EA polygenic score.

The greater average incomes and more efficient investments of agents with higher EA scores afford them higher consumption throughout their lives. Differences in

[15]Our specification of the utility cost (Equation 4) is age-dependent. We base our calculations on its values at age 50.

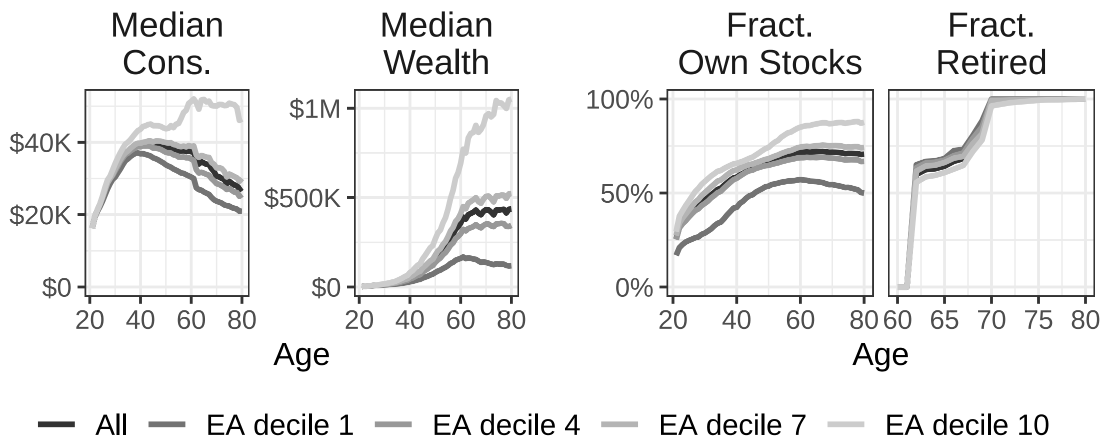
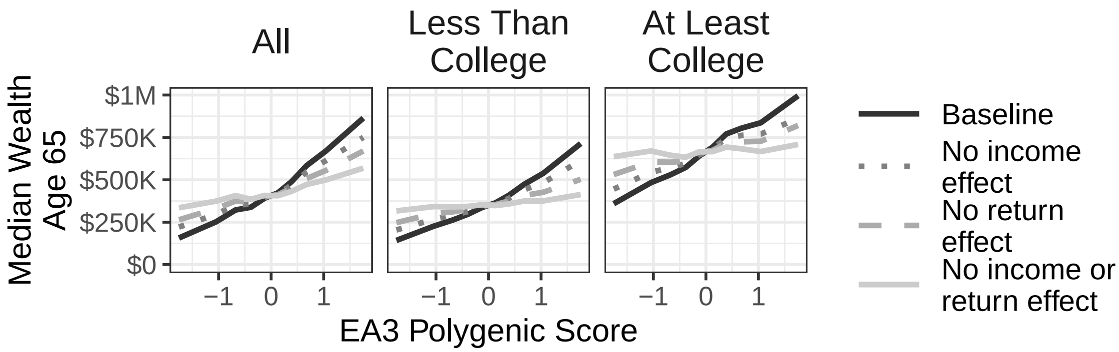


Figure 3: The choices and economic outcomes of agents with different EA scores.

median consumption become noticeable around age 35. By age 40, the difference between the median consumption of agents in the 4th and 7th deciles of the EA score reaches $970 per year; this gap grows to $3,490 by age 60 and to $4,430 by age 80.

Despite higher consumption, agents with higher EA scores accumulate more wealth. Figure 3 shows that an EA score-wealth relationship emerges early in life and widens as agents age. The median wealth balances of agents in the 4th and 7th decile of the EA polygenic score are $45,200 and $56,900 at age 40, $298,000 and $410,600 at age 60, and $344,400 and $523,100 at age 80, respectively. The growth in wealth differences is due to the combination of higher labor earnings and higher returns on savings, both of which are positively correlated with the EA score. That is, the EA score serves as a source of commonality across multiple traits that are advantageous for wealth accumulation, further magnifying differences in wealth compared to a setting in which labor earnings and returns to wealth were uncorrelated. Our results suggest that this positive correlation is partly due to genetic endowments.

Figure 3 also shows that stock market participation is monotonically increasing in the EA score, peaking at age 67 for those with EA scores in the 4th and 7th deciles, reaching values of 69% and 75%, respectively. However, this effect is due entirely to heterogeneity in returns. One common hypothesis for differences in stock market participation is that costs of participation vary widely in the population (Vissing-Jorgensen (2002), Haliassos and Michaelides (2003), Gomes and Michaelides (2005)). Instead, we find that despite virtually identical participation costs, lower EA score agents are less likely to own stocks because their expected, risk-adjusted returns are notably lower, leading to an endogenously lower participation rate. This is similar to the findings in Calvet, Campbell, and Sodini (2007).



"Baseline" corresponds to conditional medians calculated using the simulations of our estimated model. "No income effect" corresponds to simulations from a version of the model in which we set the EA score's coefficients in our income specification to 0, leaving the other components of the model unchanged. "No return effects" sets the EA score's coefficient on expected risky-asset log-returns to 0. "No income or return effect" sets both sets of coefficients to 0.

Figure 4: The EA score-wealth gradient and its sources.

Despite the lower levels of wealth and similar levels of labor disutility, differences in the retirement decisions of agents with different EA scores are small. Agents with higher EA scores retire at later ages on average. Our model matches this fact: 63% of agents in the 4th decile of the EA score retire at the minimum age of 62, and 71% have retired by age 67; for the 7th decile of the EA score, these numbers are 62% and 69%, respectively. There are a few possibilities that could explain this finding. First, lower EA score individuals will receive retirement income that represents a higher fraction of their lifetime earnings due to the progressivity of the social security system. Second, the labor earnings lost in retirement are lower for low EA score agents because they earn less income in the labor market on average. These two effects seemingly dominate the lower levels of wealth and consumption in retirement also experienced by low EA score agents.

4.3 Decomposition of the Wealth Gradient

The foregoing discussion highlights how the estimated effects of the EA score on labor income and financial proficiency compound to generate differences in wealth, stock market participation, and retirement. Among these outcomes, wealth has received the greatest attention, as a vast literature in economics has worked to understand key sources of wealth disparities. Barth, Papageorge, and Thom (2020) demonstrate that there is a robust association between the EA score and wealth at retirement. Our model reproduces this feature: the left panel of Figure 4 shows that, in the simulated populations at age 65, the difference in median wealth between agents in the 1st and

10th deciles of the EA score is is $707,400. This section quantifies the contribution of various channels to this difference.

A first driver of the relationship between the EA score and wealth is education, which the EA score was built to predict. As the middle and right panels of Figure 4 show, the relationship persists after separating agents by their level of education. The differences in median wealth at age 65 between the 1st and 10th EA score deciles fall only to $571,300 for those without a college degree and to $636,500 for those with a college degree. Therefore, the relationship between the EA score and education—and the effect of education on income, returns and other model fundamentals—explains only a small fraction of its relationship with wealth.

To evaluate the contribution of mechanisms that do not operate through education, we predict what the wealth gradients would be in counterfactual scenarios that remove different estimated effects of the EA score. Figure 4 displays these counterfactual gradients, comparing them with the baseline scenario. We first remove the estimated effects of the EA score on labor income that do not operate through its correlation with education and SES. To do so, we set the all coefficients that multiply the EA score to 0 in the income process, leaving the rest unchanged. The resulting wealth gradients are labeled "No income effect." The gene-wealth relationships flatten in this scenario. The differences in median wealth between agents in the highest and lowest EA score deciles fall to $409,500 for those without a college degree and $416,300 for college graduates, a reduction of 28% and 35%, respectively. The unconditional gradient falls by 25% to $529,100.

Next, we shut down the effect on risky asset returns by setting $r_{\text{EA}} = 0$ in Equation 6. This is produces the gradients labeled "'No return effect" in Figure 4. The effect on the gene-wealth gradient from eliminating return effects is much stronger than the effect of eliminating labor income effects, highlighting the particular importance of financial proficiency. The differences in median wealth fall by 55% to $255,600 for those without a college degree and by 55%, to $289,600, for college graduates, a significantly greater reduction than the results from removing the direct income effect. Shutting down the association between the EA score and returns flattens the unconditional gene-wealth gradient by 43%, reducing it to $405,800.

When we remove the direct effects of the EA score on both income and returns, the gaps in median wealth between the top and bottom EA deciles fall to $96,600 for those without a college degree and $72,000 for college graduates, just 17% and 11% of the baseline differences. The unconditional gap falls to $233,500, which is only one third of the initial difference. This emphasizes how critical the labor income and risky asset return channels are for establishing the link between the EA score and wealth. It also shows that the direct effects that we estimate—the ones that do not operate through education and SES—explain the vast majority of this relationship.

The small wealth differences that remain after removing the effects on income and returns are due to the combination of other direct effects, such as the EA score effect on the disutility of work and participation costs, as well as indirect effects, including the correlation of the EA score and childhood socioeconomic status, the influence of the EA score on the inheritance process, and covariation between the EA score and other characteristics such as birth year and defined benefit pension arrangements.

5 Policy Experiments

The previous section demonstrated that genetic endowments are associated with wealth accumulation primarily through risky asset returns and labor market earnings. Crucially, the extent to which these differences affect individual outcomes and welfare may depend on the policy environments individuals face. Alternate policy environments may attenuate or magnify the association between genes and wealth through the endogenous decisions of individuals in those environments, despite genetic endowments being fixed at conception. To assess the sensitivity of the gene-wealth gradient to alternative policy regimes, this section considers two counterfactual reforms that affect the quantity of available financial resources in retirement. Specifically, we consider changes to the Social Security system in the United States. Social Security constitutes a significant source of household retirement wealth (Diamond, 2004; Social Security Administration, 2016), and changes to the scale of those benefits could alter many household decisions over the lifecycle. We examine how such a change would affect the behavior, outcomes, and welfare of households with different genetic endowments, as well as the aggregate association between genetic endowments and wealth. We emphasize that this analysis would not be feasible with the available methods used in the literature on $G \times E$ interactions, which rely on *observed*, rather than counterfactual, changes in the relevant environments.

We consider changes to the Social Security system because it is one of the most consequential programs for elderly adults in the United States, and because the past few decades have seen proposals to reform this system in response to the fiscal challenges raised by an aging population. Multiple proposals have been presented every year.[16] To reduce the cost of the Social Security system, a policy would have to collect more resources from young workers, reduce the benefits distributed to retirees, or involve some combination of the two. Such policies would imply a significant reduction of households' expected lifetime financial resources, making changes in the generosity of the social safety net a natural laboratory to study effects on the link between the EA score, economic outcomes, and welfare.

[16] The Office of the Chief Actuary of the Social Security Administration maintains a repository of fiscal analyses of the proposals presented by Congress (Social Security Administration, 2022a).

We evaluate two policies. The first policy, which we call the "benefit shift," is a five-year shift in the Social Security benefit schedule. The minimum benefit-claiming age increases from 62 to 67, and the benefit calculation formulas move forward by five years. The second policy, which we call the "benefit reduction," is a simple reduction of all benefits by a constant fraction ψ. If agent i retiring at age a currently receives yearly benefits equal to $\mathrm{SSB}_i(a)$, he would receive $\mathrm{SSB}_i(a-5)$ under the benefit shift and $(1-\psi) \times \mathrm{SSB}_i(a)$ under the benefit reduction. We calibrate ψ so that both policies increase revenue by the same amount.[17] We assume agents are aware of the policy changes starting at age 21 (when the model begins), and we do not model uncertainty about policy regimes.

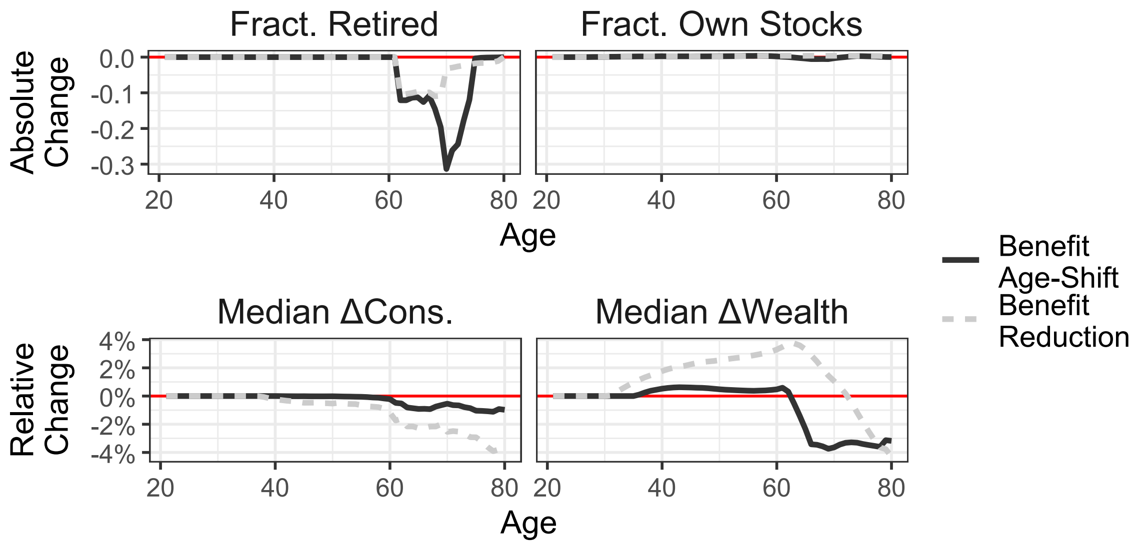
While the two policies generate the same aggregate revenue, they differ in the margins along which they incentivize individuals to adjust because they do not produce equivalent schedules of benefits. Individuals are heterogeneous across a variety of characteristics, and this induces heterogeneity in behavioral responses. For instance, individuals with lower utility costs of work may find it less costly to work for longer, while those that earn higher returns on savings may choose to cut consumption early in life. These features suggest not only different average responses, but also different distributional consequences.

5.1 Behavioral Responses

Figure 5 shows the differences in lifecycle choices under both policies relative to the baseline model. Households choose to work for longer and save more, but the magnitude of these responses differs across policies. Figure 5 shows that the benefit-shift policy causes greater delays of retirement and the benefit-reduction policy causes a greater increase in saving. For example, the fraction of households that have retired by age 70 decreases by 31 percentage points under the benefit-shift policy and only by 3 percentage points under the benefit-reduction policy. Saving follows the opposite pattern: at age 70, the median household reduces their consumption by 2.5% under the benefit-reduction policy and by only 0.5% under the benefit-shift policy.

As shown in Figure 5, wealth dynamics also change. Under the benefit-reduction policy, households build up substantially larger wealth balances prior to retirement in anticipation of lower permanent social security payments, then spend down that additional wealth in retirement. Despite the growth in pre-retirement wealth, under the benefit-reduction policy households have lower average wealth by their late 70's compared to the baseline. Under the benefit-shift policy, the difference in pre-retirement wealth and the reduction in consumption relative to the baseline are smaller. In-

[17] This calculation accounts for endogenous behavioral responses.Details are found in Appendix G. We find that the reduction that makes the two policies fiscally equivalent is $\psi = 0.2789$.



We simulate identical populations of agents that experience identical sequences of shocks under the status quo and the two alternative social security changes. We find the relative consumption and wealth changes with respect to the baseline scenario agent-by-agent and report the medians of those changes at every age. We also find the fraction of agents who are retired and who are holding stocks at every age and report the differences with respect to the baseline scenario.

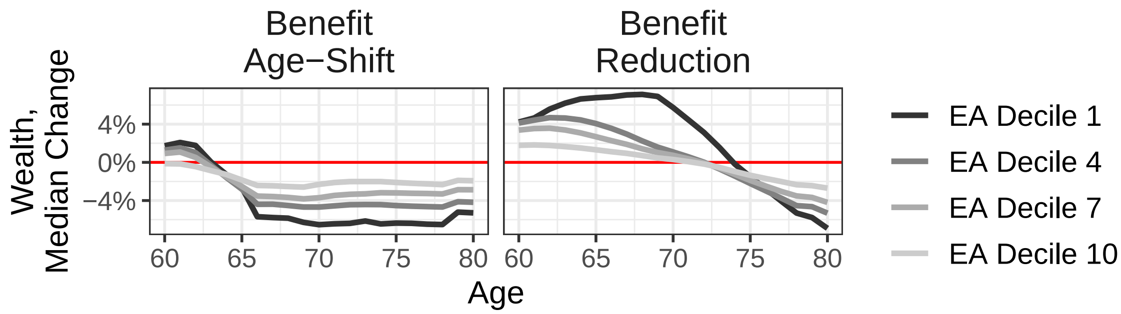
Figure 5: Life-cycle adjustments to Social Security changes.

stead, this policy inspires a greater adjustment along the dimension of labor supply, with households preferring to extend their working lives rather than dramatically increasing saving. Nonetheless, a fraction of households decide to retire prior to being eligible for Social Security benefits under the benefit-shift policy, and this rapidly reduces their wealth after age 62. In general, both policies incentivize households to work for longer and save more, but the optimal mix of these adjustments is different across policies.

Stock market participation is largely unaffected at the extensive margin. This suggests that neither policy fundamentally alters the relative attractiveness of risky assets conditional on savings decisions.

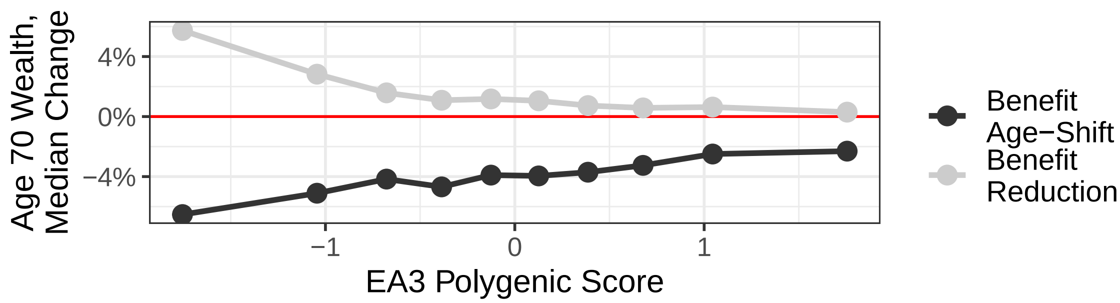
5.2 Gene-Outcome and Gene-Welfare Gradients

Next, we turn to our primary focus of the effects of these counterfactual reforms on the association between the EA score and retirement wealth. Figure 6 plots the percentage change in wealth at different ages for different levels of the polygenic score. The magnitude of the change in wealth under each policy is largest for individuals with lower polygenic scores. At age 70, the reduction in retirement wealth caused



We simulate identical populations of agents that experience identical sequences of shocks under the status quo and the two alternative Social Security changes. We find the relative wealth changes with respect to the baseline scenario agent-by-agent and report the medians of those changes at every age for different levels of the EA score.

Figure 6: Wealth response to policy changes by EA polygenic score.



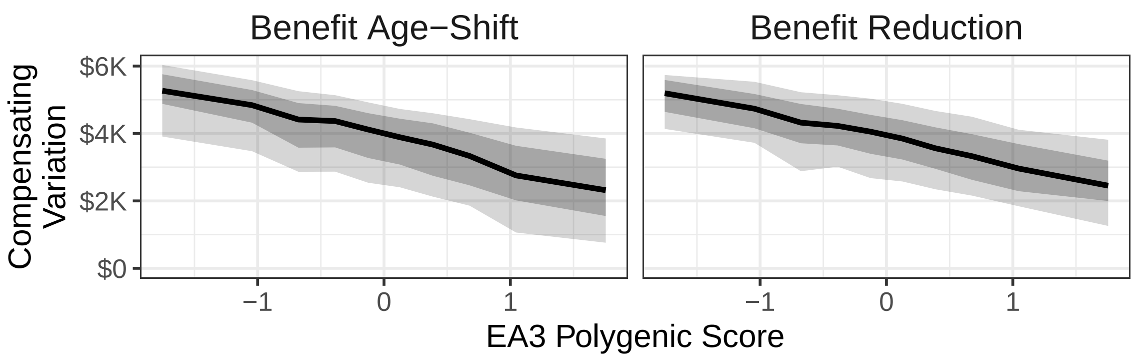
We simulate identical populations of agents that experience identical sequences of shocks under the status quo and the two alternative Social Security changes. We find the relative wealth changes at age 70 with respect to the baseline scenario agent-by-agent and report the medians of those changes for different levels of the EA score.

Figure 7: Social Security changes and the EA-wealth gradient.

by the benefit-shift policy for individuals in the lowest EA score decile (median of −6.5%) is more than double the reduction for those in the highest EA score decile (median of −2.3%). By this same age, the increases in retirement wealth caused by the benefit-reduction policy are much larger for those in the lowest EA score decile (median of +5.7%) than for those in the highest decile (median below +1%).

These differential responses alter the aggregate relationship between the polygenic score and retirement wealth. Figure 7 plots the relative changes in wealth under both policies compared to the baseline for individuals at age 70 against the polygenic score. Under the benefit-reduction policy the gene-wealth gradient actually flattens, as lower polygenic score households accumulate relatively more wealth in anticipation of reduced retirement benefits than do higher polygenic score households. Under the benefit-shift policy the gene-wealth gradient steepens, with lower polygenic score households experiencing a larger decline in financial resources at age 70.Despite the fiscal equivalence of the two policies, the implications for the slopes of the gene-wealth gradients are notably different between them.

To what extent do changes in the relationship between genes and wealth accurately proxy for the change in the relationship between genes and well-being? To answer this question, we use our model to calculate compensating variations, the monetary amounts by which individuals would need to be compensated to preserve their welfare after the policy reforms.[18] An individual will require a greater compensating variation for a policy that reduces his welfare by a greater amount. We calculate compensating variations from the point of view of individuals who are informed of the policy changes at age 21 and who have a starting wealth of $20,000.[19]



The compensating variations for both policies are computed at age 21 and from a starting wealth of $20,000. We find the compensating variation for each of our $5,130$ agent types and weight them by the number of agents of each type in our simulated population. The figure depicts percentiles of the distribution of these compensating variations at different levels of the EA score. The solid line corresponds to the median. Inner shaded areas cover observations between the 25th and 75th percentiles. Outer shaded areas cover observations between the 10th and 90th percentile.

Figure 8: Distribution of compensating variations at different levels of the EA score.

Figure 8 shows that, despite their different implications for the gene-wealth gradient, both policy changes steepen the gene-*welfare* gradient, as each generates greater welfare shortfalls for individuals with lower polygenic scores. The unconditional distribution of the compensating variations of both policies is similar: the mean and standard deviation for the benefit-shift policy are $3,800 and $1,300 respectively, and

[18]If $V_{i,21}(\cdot)$ is the baseline value function of agent i at age 21 and $V^x_{i,21}(\cdot)$ is his value function under the alternative policy scenario x, the compensating variation from an initial wealth of W, $\mathrm{CV}^x_i(W)$, solves $V_{i,21}(W) = V^x_{i,21}\left(W + \mathrm{CV}^x_i\left(W\right)\right)$. The compensating variation for a given policy is a function of the agent's characteristics—expected income path, preferences, financial proficiency, etc.—and his wealth W.

[19]The distribution of individuals across characteristics matches that of the sample that we use for estimating the model.

for the benefit-reduction policy, they are \$3,800 and \$1,100. However, within both policy regimes, there are large differences in the welfare costs borne by agents with different polygenic scores. In both cases, the median compensating variation of individuals in the lowest decile of the polygenic score is more than twice as large as that of agents in the highest decile of the polygenic score.

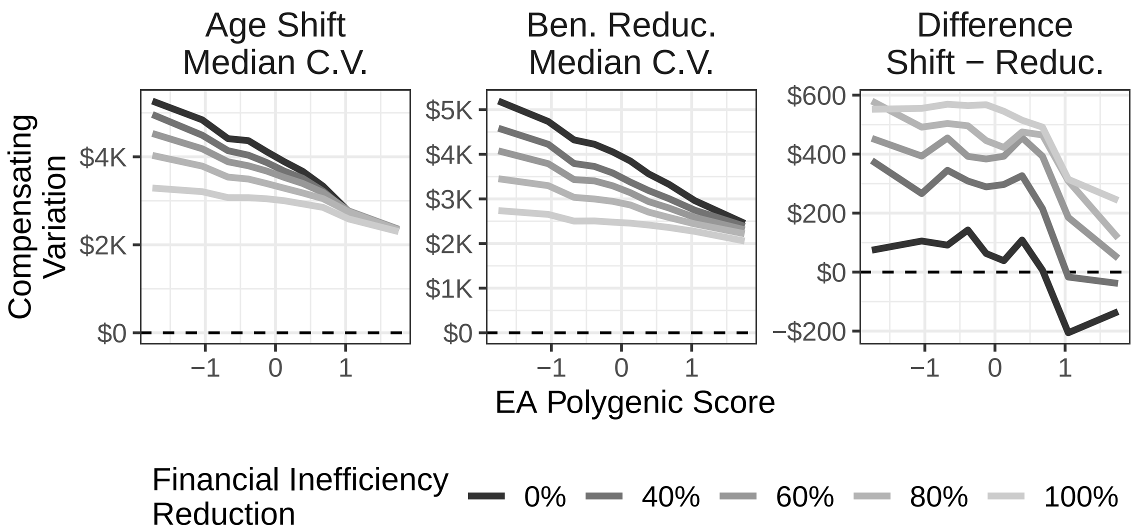
We highlight that under the benefit-reduction policy, the gene-wealth gradient flattens, suggesting genetic inequality has *lessened* under this policy. Yet, under this same policy the gene-*welfare* gradient actually steepens, indicating that despite smaller differences in wealth, differences in well-being are magnified as lower polygenic score individuals are especially harmed. This exercise demonstrates that gene-outcome associations are both policy-environment sensitive and may be a poor representation of the underlying association between genetic endowments and welfare. Comparing the effects of different policies based on the relationship between genes and outcomes may obscure their effects on the ultimate quantity of interest, welfare.

5.3 Investment Efficiency and Welfare Changes

Changes in the generosity of retirement benefits may heighten the importance of sound financial decisions. In this section, we examine the extent to which heterogeneity in financial proficiency drives the differential impact of the policy reforms, particularly between individuals with different polygenic scores.

We consider a scenario in which an intervention improves the financial proficiency of our agents, lowering their participation costs and increasing the efficiency of their stock investments. Determining the nature of such a policy is well outside the scope of this paper; we refer the reader to the growing literature that evaluates the effectiveness of different types of interventions aimed at improving financial knowledge and behaviors (see Kaiser and Menkhoff, 2017; Kaiser et al., 2022). Rather than choosing a level of effectiveness for this hypothetical intervention, we consider a range of efficiency improvements: 40%, 60%, 80%, and 100%. We reduce the stock market participation cost and the risky asset return inefficiency of every agent by each level. For instance, under an efficiency improvement of 40%, we reduce every household's participation cost by 40% and move every household 40% closer to the stock market's benchmark expected log-return than they were under the baseline estimates.

Figure 9 shows how the gene-welfare gradients of both policies would change with improvements to financial efficiency. Heterogeneous inefficiencies explain a large part of the gradient: as the population becomes more financially efficient, the gradients flatten. With a 100% reduction of financial inefficiencies, the slope of the gene-welfare gradient would be 66% smaller for the benefit-shift policy and 75% smaller for the benefit-reduction policy. Changes in financial proficiency also alter the relative ranking of the two policies. As financial inefficiencies fall, the compensating variation



The compensating variations for both policies are computed at age 21 and from a starting wealth of $20,000. We find the compensating variation for each of our $5,130$ agent types and weight them by the number of agents of each type in our simulated population. For both policies and each level of financial inefficiency reduction, we find the median compensating variation at different levels of the EA score. These are presented in the left and middle panels. Each point in the right panel subtracts the median compensating variation of the benefit reduction policy (at the given inefficiency reduction and EA score) from that of the age-shift policy.

Figure 9: Reductions of Financial Inefficiencies and Welfare Costs.

in the benefit-shift policy becomes significantly larger than in the benefit-reduction policy. This is due to the different adjustments that the policies induce; since the adjustments under the benefit-reduction policy rely more heavily on an individual's savings, improving their management of these savings makes the policy more attractive than the alternative. These findings highlight that an intervention that reduces heterogeneity in a certain dimension of wealth accumulation can alter the distributions of welfare consequences of various policies across multiple characteristics, including genes.

Our results demonstrate that policies can shift $G \times E$ interactions in complex ways. Not only does the policy environment affect the gradients between observed economic outcomes and genetic endowments, but different policies also lead to different distributions of welfare costs for people with different genetic endowments. We highlight that, absent a model that explicitly incorporates both genetic endowments and endogenous behavioral responses to changing environments, such conclusions would not be possible.

6 Conclusion

We allow genetic endowments associated with educational attainment, summarized as an index known as a polygenic score, to directly affect multiple dimensions of wealth accumulation in a structural model of lifecycle consumption and savings. We show that the polygenic score affects wealth primarily through labor market earnings and the return on risky investments. Heterogeneity in returns appears to be particularly important. Because genes are passed from parents to offspring and fixed at conception, they represent one potential source of the persistence in returns over time and across generations. We further show that in a counterfactual setting with less generous public pension benefits, the gene-wealth gradient flattens, suggesting genetic inequality in wealth has weakened, but the gene-wealth gradient actually steepens. This highlights that reduced-form associations between genes and economic outcomes may be an imperfect representation of their underlying associations with well-being.

The positive association between the polygenic score for education and multiple factors that promote wealth accumulation magnifies the inequality in wealth. High polygenic score individuals are more likely to be college graduates, earn more in the labor market conditional on education, and earn higher returns on savings. Unlike patience for instance, which could affect labor market earnings through years of schooling but would have no direct implication for earnings conditional on education, or risk tolerance, which may associate with riskier and higher-paying careers and a greater allocation to risky assets but have no consequence for the returns earned on risky assets, the genetic endowments we study represent a single factor that promotes wealth accumulation through multiple complementary channels.

Further, because genes are inherited, they constitute one potential source for the intergenerational persistence in wealth. Even in the absence of direct transfers via monetary inheritances, the transmission of traits that are advantageous for wealth accumulation would create a positive correlation in wealth across generations, and our results suggest the channels through which these correlations may operate. In fact, ignoring this transmission of advantageous traits may have consequences for predictions of the effectiveness of policies intended to improve socioeconomic mobility.

This paper contributes to the growing literature on biology and heterogeneity in wealth by uncovering the underlying mechanisms that link particular genetic endowments to wealth, and provides a roadmap for including genetic variation in structural models of optimizing economic behavior in future work.

References

A. Adermon, M. Lindahl and D. Waldenström. 2018. “Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations.” *The Economic Journal* 128 (612):F482–F513.

B. W. Arold, P. Hufe and M. Stoeckli. 2022. “Genetic Endowments, Educational Outcomes and the Mediating Influence of School Investments.” *CESifo Working Papers* (9841).

O. Attanasio, H. Low and V. Sánchez-Marcos. 2008. “Explaining Changes in Female Labor Supply in a Life-Cycle Model.” *American Economic Review* 98 (4):1517–1552.

G. Auten and D. Splinter. 2024. “Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends.” *Journal of Political Economy* 132 (7):2179–2227.

L. Bach, L. E. Calvet and P. Sodini. 2020. “Rich Pickings? Risk, Return, and Skill in Household Wealth.” *American Economic Review* 110 (9):2703–2747.

C. Badarinza, J. Y. Campbell and T. Ramadorai. 2016. “International Comparative Household Finance.” *Annual Review of Economics* 8 (Volume 8, 2016):111–144.

S. Bagchi. 2015. “Labor Supply and the Optimality of Social Security.” *Journal of Economic Dynamics and Control* 58:167–185.

S. Baker, P. Biroli, H. van Kippersluis et al. 2022. “Beyond Barker: Infant Mortality at Birth and Ischaemic Heart Disease in Older Age.”

S. H. Barcellos, L. Carvalho and P. Turley. 2021. “The Effect of Education on the Relationship between Genetics, Early-Life Disadvantages, and Later-Life SES.”

A. Barnea, H. Cronqvist and S. Siegel. 2010. “Nature or Nurture: What Determines Investor Behavior?” *Journal of Financial Economics* 98 (3):583–604.

D. Barth, N. W. Papageorge and K. Thom. 2020. “Genetic Endowments and Wealth Inequality.” *Journal of Political Economy* 128 (4):1474–1522.

J. Becker, C. A. P. Burik, G. Goldman et al. 2021. “Resource Profile and User Guide of the Polygenic Index Repository.” *Nature Human Behaviour* 5 (12):1744–1758.

D. W. Belsky, B. W. Domingue, R. Wedow et al. 2018. “Genetic Analysis of Social-Class Mobility in Five Longitudinal Studies.” *Proceedings of the National Academy of Sciences* 115 (31):E7275–E7284.

D. W. Belsky, T. E. Moffitt, D. L. Corcoran et al. 2016. “The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development.” *Psychological Science* 27 (7):957–972.

J. Benhabib, A. Bisin and M. Luo. 2019. “Wealth Distribution and Social Mobility in the US: A Quantitative Approach.” *American Economic Review* 109 (5):1623–1647.

J. Benhabib, A. Bisin and S. Zhu. 2011. “The Distribution of Wealth and Fiscal

Policy in Economies With Finitely Lived Agents." *Econometrica* 79 (1):123–157.

L. Bierut, P. Biroli, T. J. Galama et al. 2023. "Challenges in studying the interplay of genes and environment: A study of childhood financial distress moderating genetic predisposition for peak smoking." Working paper.

P. Biroli, T. J. Galama, S. von Hinke et al. 2022. "The Economics and Econometrics of Gene-Environment Interplay."

S. E. Black, P. J. Devereux, P. Lundborg et al. 2017. "On the Origins of Risk-Taking in Financial Markets." *The Journal of Finance* 72 (5):2229–2278.

———. 2020. "Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviours." *The Review of Economic Studies* 87 (4):1683–1725.

R. Blundell, M. Costa Dias, C. Meghir et al. 2016. "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica* 84 (5):1705–1753.

L. E. Calvet, J. Y. Campbell and P. Sodini. 2007. "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes." *Journal of Political Economy* 115 (5):707–747.

J. Y. Campbell, T. Ramadorai and B. Ranish. 2019. "Do the Rich Get Richer in the Stock Market? Evidence from India." *American Economic Review: Insights* 1 (2):225–240.

C. D. Carroll. 2002. "Why Do the Rich Save So Much?" In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, edited by Joel B. Slemrod. Harvard University Press.

S. Catherine. 2021. "Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle." *The Review of Financial Studies* :hhab136.

S. Catherine, M. Miller and N. Sarin. 2020. "Social Security and Trends in Wealth Inequality."

D. Cesarini, C. T. Dawes, M. Johannesson et al. 2009. "Genetic Variation in Preferences for Giving and Risk Taking*." *The Quarterly Journal of Economics* 124 (2):809–842.

D. Cesarini, M. Johannesson, P. Lichtenstein et al. 2010. "Genetic Variation in Financial Decision-Making." *The Journal of Finance* 65 (5):1725–1754.

J. Chai, W. Horneff, R. Maurer et al. 2011. "Optimal Portfolio Choice over the Life Cycle with Flexible Work, Endogenous Retirement, and Lifetime Payouts*." *Review of Finance* 15 (4):875–907.

J. F. Cocco, F. Gomes and P. J. Maenhout. 2005. "Consumption and Portfolio Choice over the Life Cycle." *The Review of Financial Studies* 18 (2):491–533.

M. D. Collado, I. Ortuño-Ortín and J. Stuhler. 2023. "Estimating Intergenerational and Assortative Processes in Extended Family Data." *The Review of Economic Studies* 90 (3):1195–1227.

H. Cronqvist and S. Siegel. 2014. “The Genetics of Investment Biases.” *Journal of Financial Economics* 113 (2):215–234.

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

M. De Nardi, E. French and J. B. Jones. 2010. “Why Do the Elderly Save? The Role of Medical Expenses.” *Journal of Political Economy* 118 (1):39–75.

P. Diamond. 2004. “Social Security.” *American Economic Review* 94 (1):1–24.

K. E. Dynan, J. Skinner and S. P. Zeldes. 2004. “Do the Rich Save More?” *Journal of Political Economy* 112 (2):397–444.

A. Fagereng, C. Gottlieb and L. Guiso. 2017. “Asset Market Participation and Portfolio Choice over the Life-Cycle.” *The Journal of Finance* 72 (2):705–750.

A. Fagereng, L. Guiso, D. Malacrino et al. 2020. “Heterogeneity and Persistence in Returns to Wealth.” *Econometrica* 88 (1):115–170.

A. Fagereng, M. Mogstad and M. Rønning. 2021. “Why Do Wealthy Parents Have Wealthy Children?” *Journal of Political Economy* 129 (3):703–756.

J. M. Fletcher and Q. Lu. 2021. “Health policy and genetic endowments: Understanding sources of response to Minimum Legal Drinking Age laws.” *Health Economics* 30 (1):194–203.

S. Furuya, J. M. Fletcher, Z. Zhao et al. 2022. “Detecting genetic heterogeneities in response to trauma: The case of 9/11.” *SSM - Mental Health* 2.

H.-M. V. Gaudecker. 2015. “How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?” *The Journal of Finance* 70 (2):489–507.

G.-L. Gayle, L. Golan and M. A. Soytas. 2022. “What Is the Source of the Intergenerational Correlation in Earnings?” *Journal of Monetary Economics* 129:24–45.

F. Gomes, M. Haliassos and T. Ramadorai. 2021. “Household Finance.” *Journal of Economic Literature* 59 (3):919–1000.

F. Gomes and A. Michaelides. 2005. “Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence.” *The Journal of Finance* 60 (2):869–904.

M. Haliassos and A. Michaelides. 2003. “Portfolio Choice and Liquidity Constraints.” *International Economic Review* 44 (1):143–177.

J. Heckman. 1974. “Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption Over the Life Cycle.” *The American Economic Review* 64 (1):188–194.

W. D. Hill, N. M. Davies, S. J. Ritchie et al. 2019. “Genome-Wide Analysis Identifies Molecular Systems and 149 Genetic Loci Associated with Income.” *Nature Communications* 10 (1):5741.

T. Kaiser, A. Lusardi, L. Menkhoff et al. 2022. “Financial Education Affects Financial Knowledge and Downstream Behaviors.” *Journal of Financial Economics* 145 (2, Part A):255–272.

T. Kaiser and L. Menkhoff. 2017. "Does Financial Education Impact Financial Literacy and Financial Behavior, and If So, When?" *The World Bank Economic Review* 31 (3):611–630.

R. Karlsson Linnér, P. Biroli, E. Kong et al. 2019. "Genome-Wide Association Analyses of Risk Tolerance and Risky Behaviors in over 1 Million Individuals Identify Hundreds of Loci and Shared Genetic Influences." *Nature Genetics* 51 (2):245–257.

A. Kong, G. Thorleifsson, M. L. Frigge et al. 2018. "The Nature of Nurture: Effects of Parental Genotypes." *Science* 359 (6374):424–428.

J. J. Lee, R. Wedow, A. Okbay et al. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50 (8):1112–1121.

H. Low, C. Meghir and L. Pistaferri. 2010. "Wage Risk and Employment Risk over the Life Cycle." *American Economic Review* 100 (4):1432–1467.

A. Lusardi, P.-C. Michaud and O. S. Mitchell. 2017. "Optimal Financial Knowledge and Wealth Inequality." *Journal of Political Economy* 125 (2):431–477.

A. R. Martin, C. R. Gignoux, R. K. Walters et al. 2017. "Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations." *The American Journal of Human Genetics* 100 (4):635–649.

R. C. Merton. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case." *The Review of Economics and Statistics* 51 (3):247–257.

D. Muslimova, H. van Kippersluis, C. A. Rietveld et al. 2020. "Dynamic Complementarity in Skill Production: Evidence From Genetic Endowments and Birth Order." SSRN Scholarly Paper 3748468, Social Science Research Network, Rochester, NY.

A. Okbay, J. P. Beauchamp, M. A. Fontana et al. 2016. "Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment." *Nature* 533 (7604):539–542.

A. Okbay, Y. Wu, N. Wang et al. 2022. "Polygenic Prediction of Educational Attainment within and between Families from Genome-Wide Association Analyses in 3 Million Individuals." *Nature Genetics* 54 (4):437–449.

A. Olsen and R. Hudson. 2009. "Social Security Administration's Master Earnings File: Background Information." *Social Security Bulletin* 82 (4).

S. Ozkan, J. Hubmer, S. Salgado et al. 2023. "Why Are the Wealthiest So Wealthy? New Longitudinal Empirical Evidence and Implications for Theories of Wealth Inequality."

N. W. Papageorge and K. Thom. 2020. "Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study." *Journal of the European Economic Association* 18 (3):1351–1399.

C. A. Rietveld, S. E. Medland, J. Derringer et al. 2013. "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science*

340 (6139):1467–1471.

V. Ronda, E. Agerbo, D. Bleses et al. 2022. "Family Disadvantage, Gender, and the Returns to Genetic Human Capital*." *The Scandinavian Journal of Economics* 124 (2):550–578.

A. Rustichini, W. G. Iacono, J. J. Lee et al. 2023. "Educational Attainment and Intergenerational Mobility: A Polygenic Score Analysis." *Journal of Political Economy* 131 (10):2724–2779.

J. Sabelhaus and A. H. Volz. 2022. "Social Security Wealth, Inequality, and Life-Cycle Saving." In *Measuring Distribution and Mobility of Income and Wealth*. University of Chicago Press, 249–286.

E. Saez and G. Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data *." *The Quarterly Journal of Economics* 131 (2):519–578.

P. A. Samuelson. 1969. "Lifetime Portfolio Selection By Dynamic Stochastic Programming." *The Review of Economics and Statistics* 51 (3):239–246.

Social Security Administration. 2016. "Income of the Population 55 or Older, 2014."

———. 2022a. "Proposals to Change Social Security." https://www.ssa.gov/oact/solvency/index.html.

———. 2022b. "Summary of Provisions That Would Change the Social Security Program." Tech. rep., Social Security Administration.

P. Taubman. 1976. "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins." *The American Economic Review* 66 (5):858–870.

S. Trejo and B. W. Domingue. 2018. "Genetic Nature or Genetic Nurture? Introducing Social Genetic Parameters to Quantify Bias in Polygenic Score Analyses." *Biodemography and Social Biology* 64 (3-4):187–215.

A. Vissing-Jorgensen. 2002. "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures." Working Paper 8884, National Bureau of Economic Research.

A. I. Young, M. L. Frigge, D. F. Gudbjartsson et al. 2018. "Relatedness Disequilibrium Regression Estimates Heritability without Environmental Bias." *Nature Genetics* 50 (9):1304–1310.