Online Appendix:

"Genetic Endowments and Wealth Inequality"

By: Daniel Barth, Nicholas W. Papageorge and Kevin Thom[1]

A Wealth Data and Additional Summary Statistics

A.1 Retirement Wealth

The wealth data used in this paper are largely constructed from the RAND wealth and income files. The RAND files are carefully cleaned and consistently coded by RAND Corportation and are available for public use. The RAND files have been used in both academic and industry publications, and ensure comparability and consistency across HRS waves and research projects. We refer the reader to the RAND codebook and documentation for further details.

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

Unfortunately, data on employer-sponsored retirement plans are not asked in every wave, and are sometimes inconsistently coded across waves. The remainder of this section focuses on our methodology for coding retirement account balances and stock market participation inferred from those accounts. Broadly speaking, there are two types of retirement plans: defined-benefit plans, such as traditional pensions (which the HRS calls type A plans), and defined contribution plans, such as 401k and 403(b) plans (which the HRS calls type B plans). We discuss each type of plan in turn.

A.1.1 Defined Benefit Plans

To deal with issues arising from type A style retirement plans, our sample includes only households fully in retirement (households in which no member of the household is currently working). We exclude working households because expected benefits from defined-benefit

[1]This version of the online appendix was generated on July 11, 2019.

pension plans are likely to be both an important source of wealth and noisily measured. For retired households, our assumption is that those who report receiving pension income were included in defined-benefit pension plans at some point during their working lives, and those who do not receive pension income in retirement were not included in such plans. To the extent that households misreport pension income, for example if income from an annuity converted from a 401k plan is reported as pension income, or if households have delayed receiving pension benefits until some future date, our assignment of households participating in type A plans will be biased. Further, because the household earns a guaranteed stream of income regardless of the underlying investments that support that income (and because we do not observe these underlying investments), we do not consider a household's participation in type A pension plans to be participation in the stock market.

We include retirement income in our household wealth measure by calculating the price of an actuarially fair annuity based on the entirety of household retirement income, which includes pension income, annuity income, and income from social security. We follow Yogo (2016) by calculating the present discounted value of this income based on a 1.5% annual risk-free rate of return, and discount income in each year by the probability of the recipient surviving until that year.[2] Specifically, we calculate the present value of retirement income, P_t, as:

$$P_t = Y_t \sum_{s=1}^{T-t-1} \left[\prod_{u=1}^{s} p_{t+u} R^{-1} \right], \tag{1}$$

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

A.1.2 Defined Contribution Plans

Wealth in defined contribution style plans is a bit trickier. Households may have plans associated with multiple previous employers. To calculate comprehensive measures of wealth and stock market participation, we would like to know both the balances and asset allocations of all employer-sponsored type B plans from all previous jobs. Unfortunately, this is not always possible.

In years 1996, 1998, and 2002-2010 (comprising even-numbered years), we have the highest quality data on total balances in employer-sponsored type B retirement plans.[3] In these

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

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

years, our wealth data include balances of employer-sponsored plans that are still maintained through that employer, and have not been converted to annuities or rolled over into IRAs. The HRS refers to such plans as *dormant plans*. Unfortunately, the value of dormant plans at employers prior to retirement are not asked in 1992, 1994, and 2000.

Dormant plans also present problems for measurement of stock market participation. While in years 2002-2010 the stock allocation within a respondent's retirement plan at the current employer is observable for working households, the stock allocation in dormant plans for retired households is not. Thus, for retirement plans at current or former employers, only stock market participation for plans at the current employer are included in our measure of stock market participation. Because our sample includes only retired households, stock market allocations at employer-sponsored plans contribute negligibly to our stock market ownership variable.

A.2 Additional Summary Statistics

This appendix provides additional summary statistics for the genotyped versus non-genotyped samples, demographics by household structure, income and wealth, and the potential mechanisms studied in Section 5. Table S1 documents differences in birth year, education, and other variables between the genotyped and non-genotyped samples. Genotyped individuals tend to be older, more educated, and belong to wealthier households. Table S2 presents the mean and standard deviation of demographic variables such as birth year, years of schooling, and highest degree earned. These are reported for all households in our sample and separately by household structure.

Panel A of Table S3 shows the distribution of lifetime labor earnings for all households and by household structure: coupled, male only, and female only households. Table S4 shows the distribution of various measures of household wealth: total wealth, total wealth excluding housing wealth, total wealth excluding the present value of retirement income, and total wealth excluding both retirement wealth and housing. Table S5 presents the mean, median, 75^{th} percentile, and 90^{th} percentile of each individual component of household wealth for the full sample for all household-years. We also calculate the share of total real wealth in each component for each household-year, and present the median and mean values of these shares.

Table S6 shows summary statistics for the various mechanisms that may help to explain the gene-wealth gradient. Panel A presents statistics on the fraction of households that receive an inheritance, the (total) amount of inheritances received for households that receive them, and the household average of parental education (years of schooling). Panel B reports

summary statistics related to mortality. In particular, we examine the annual death rate for individuals in our main sample (excluding years before genotyping), and the subjective probability of living until the age 75, which is elicited multiple times per individual. Panel C reports summary statistics for measures of risk aversion. Specifically, Panel C shows the fraction of households that fall within each risk aversion bin (based on the gambles they would accept and reject) for the labor income gamble, the business wealth gamble, and the inheritance gamble. Panel D reports the mean and standard deviation of home ownership, business ownership, and stock market participation. Stock market participation includes equity investments in retirement accounts, mutual funds, or in individually held stocks.

Panel E of Table S6 reports summary statistics related to beliefs about macroeconomic events: the probability the stock market appreciated over the following year, the deviation of this belief from the objective estimate based on historical data, and whether the respondent reports a belief of 0% or 100% (extreme beliefs), or of 50% (a focal belief that may represent uncertainty). Similar statistics are reported for the likelihood of a major (economic) depression and double-digit inflation. Panel E also reports the distribution of the reported financial planning horizon. Finally, panel F reports the fraction of households that receive defined-benefit pension income, and the present discounted value of lifetime defined benefit income.

Appendix Table S1: Selection into Genotyping

	Genotyped	Non-Genotyped	Δ p-value
Birth Year	1938.39	1941.39	0.00
Education	12.58	11.94	0.00
Male	0.41	0.45	0.00
Total Income (in $1000)	1,076.69	841.67	0.00
Wealth (in $1000)	722.90	457.54	0.00
N (Max)	12,505	30,699	

Notes: This table provides summary statistics for the genotyped and the non-genotyped individuals in the HRS data. Wealth is measured once per individual when the individual's household is first observed.

Appendix Table S2: Summary Statistics — Birth Year and Education

	All Households (2369)		Female Only (541)		Male Only (187)	
	Mean	SD	Mean	SD	Mean	SD
	[1]	[2]	[3]	[4]	[5]	[6]
Year of birth						
Female	1935.10	5.59	1933.96	5.85	.	.
Male	1933.04	5.76	.	.	1934.83	5.56
Years of Education						
Female	12.67	2.30	12.65	2.45	.	.
Male	12.74	2.96	.	.	12.56	3.18
No Degree						
Female	0.16	0.37	0.19	0.39	.	.
Male	0.19	0.39	.	.	0.17	0.38
GED						
Female	0.04	0.19	0.03	0.17	.	.
Male	0.06	0.24	.	.	0.11	0.31
High School						
Female	0.60	0.49	0.58	0.49	.	.
Male	0.47	0.50	.	.	0.46	0.50
4-Year Degree						
Female	0.10	0.30	0.08	0.27	.	.
Male	0.13	0.34	.	.	0.12	0.33
Masters						
Female	0.05	0.22	0.07	0.25	.	.
Male	0.08	0.27	.	.	0.06	0.25
Professional						
Female	0.01	0.09	0.02	0.13	.	.
Male	0.03	0.18	.	.	0.03	0.16

Notes: This table reports means and standard deviations (SD) for birth year, years of education, and indicator variables for highest degree obtained for household members examined in our main analytic sample. Summary statistics are presented for males and females in all households and then separately for female-only households and male-only households.

Appendix Table S3: Summary Statistics — Income

	Quantiles of Income (in $1000)							
Panel A	10	25	50	75	90	Mean	SD	N
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
All Households	569.86	1287.80	2255.30	3082.30	4034.19	2315.95	1405.43	2377
Coupled	941.32	1707.75	2546.66	3340.43	4272.74	2609.35	1362.69	1702
Female Only	192.82	565.01	1212.20	2029.72	2870.63	1435.51	1133.95	503
Male Only	429.36	1123.86	1767.46	2538.92	3584.53	1987.43	1397.68	172

	All Households		Coupled		Female Only		Male Only	
Panel B	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Non-Missing Income	0.93	0.26	0.93	0.25	0.93	0.26	0.92	0.27
Zero Income	0.00	0.05	0.00	0.00	0.01	0.10	0.01	0.08
Avg. Yrs. Top-Coded	12.67	12.55	15.58	12.47	3.50	7.72	10.72	11.40
Never Top Coded	0.27	0.44	0.15	0.36	0.68	0.47	0.22	0.42

Notes: This table reports summary statistics for the income measure used in our main analysis obtained from the Social Security Administration Master Earnings File data. Panel A provides information on the proportion of households with non-missing data, the proportion of households with zero income (conditional on having non-missing income data), average number of years that any household member reports top-coded income, the fraction of households never observed with a top-coded income observation, and average total real income (in $1,000s). In years in which a household member's income is top-coded, we replace the top-coded amount with the average of individual earned incomes greater than or equal to the top-coded amount in the Current Population Survey for that year. Income statistics are provided for all households and separately by household structure. Panel B reports quantiles of the income distribution (in $1000s) along with the mean and standard deviation of income by household structure.

Appendix Table S4: Summary Statistics — Wealth (in $1000)

	p10	p25	p50	p75	p90	Mean	SD
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
TotalWealth	168.74	303.82	593.64	1031.48	1706.83	838.05	851.27
NoHousing	146.12	236.99	450.49	815.63	1373.81	727.80	1286.55
NoRet	1.39	75.08	235.98	574.86	1164.70	530.36	1236.09
NoHouseorRet	0.00	6.08	92.00	339.28	838.10	357.99	1105.99

Notes: This table reports summary statistics on the distribution of real wealth in our main sample. Measures of wealth include total real wealth, total wealth excluding housing, total wealth excluding retirement wealth (the present discounted value of retirement income), and total wealth excluding housing and retirement wealth. Each wealth measure is winsorized at the 1st and 99th percentiles. These statistics are calculated for the full analytical sample of 5,701 household-year observations.

Appendix Table S5: Wealth Distribution ($1000)

	p50	p75	p90	Mean	Median Share	Mean Share
Ret Plans (Employer)	0.00	0.00	0.00	3.18	0.00	0.00
Ret Inc (PV)	278.30	436.69	664.24	365.76	0.57	0.57
Real Estate	0.00	0.00	60.77	46.98	0.00	0.03
Business	0.00	0.00	0.00	48.32	0.00	0.02
IRAs	0.00	86.95	247.49	91.69	0.00	0.07
Stocks	0.00	32.45	202.56	87.78	0.00	0.05
Cash Equiv.	8.08	28.36	75.00	30.65	0.01	0.04
CDs	0.00	3.04	53.68	20.06	0.00	0.02
Bonds	0.00	0.00	0.00	14.60	0.00	0.01
Other Assets	0.00	0.00	21.63	18.31	0.00	0.01
Other Debts	0.00	0.06	5.06	2.32	0.00	0.01
Trusts	0.00	0.00	0.00	1.91	0.00	0.00
Home Value	126.98	216.33	363.63	168.67	0.18	0.32
Mortgage	0.00	0.00	64.00	16.95	0.00	0.13
Home Loan	0.00	0.00	0.00	2.48	0.00	0.00
Second Home	0.00	0.00	40.51	24.83	0.00	0.02
Second Mortgage	0.00	0.00	0.00	1.29	0.00	0.00

Notes: Summary statistics for different sources of wealth (in $1000s). For each household-year, we calculate the share of total wealth from each source, and Columns [5] - [6] report the median and mean shares. We note that although we report positive values for Mortgages, Home Loans, and Other Debts here, these are subtracted in the construction of total wealth. Note that Ret Plans (Employer) represent only retirement accounts that are still maintained by the employer despite the household being retired.

Appendix Table S6: Summary Statistics: Mechanisms

	Mean	SD	N
Panel A: Transfers	[1]	[2]	[3]
Any Inheritance	0.41	0.49	2556
Inheritance Amount (in $1000)	160.62	612.61	1054
Fathers' Education (HH Avg.)	9.47	3.20	2294
Mothers' Educ (HH Avg.)	9.95	2.79	2345
Panel B: Mortality	[1]	[2]	[3]
Mortality	0.04	0.20	26733
Mortality Expectations	67.39	26.10	29119
Panel C: Risk Aversion	[1]	[2]	[3]
Not take 50-50 Gamble Doubling Income or 10% Cut	0.39	0.49	10512
Take 50-50 Gamble Doubling Income or 10% Cut (but not 20%)	0.22	0.42	10512
20% Cut (but not 33%)	0.17	0.37	10512
33% Cut (but not 50%)	0.10	0.30	10512
50% Cut (but not 75%)	0.07	0.26	10512
75% Cut	0.05	0.22	10512

Notes: This table reports means and standard deviations for additional variables used to investigate mechanisms underlying the estimated gene-wealth gradient. Each panel corresponds to an alternative mechanism. Mechanisms include Transfers (Panel A); Mortality (Panel B); Risk aversion (Panel C); Portfolio choices (Panel D); Beliefs and planning horizons (Panel E); and Defined-benefit pensions (Panel F). *This table continues onto the following page.*

Appendix Table S6: Summary Statistics: Mechanisms (continued)

	Mean	SD	N
Panel C: Risk Aversion (continued)	[1]	[2]	[3]
Not take 50-50 Gamble Doubling Business or 10% Cut	0.47	0.50	2912
Take 50-50 Gamble Doubling Business or			
10% Cut (but not 20%)	0.11	0.32	2912
20% Cut (but not 33%)	0.11	0.32	2912
33% Cut (but not 50%)	0.10	0.30	2912
50% Cut (but not 75%)	0.11	0.31	2912
75% Cut	0.09	0.29	2912
Not take 50-50 Gamble Doubling Inheritance or 10% Cut	0.51	0.50	2951
Take 50-50 Gamble Doubling Inheritance or			
10% Cut (but not 20%)	0.19	0.39	2951
20% Cut (but not 33%)	0.13	0.34	2951
33% Cut (but not 50%)	0.05	0.22	2951
50% Cut (but not 75%)	0.05	0.22	2951
75% Cut	0.07	0.25	2951
Panel D: Portfolio Choices	[1]	[2]	[3]
Has House	0.84	0.37	6460
Has Business	0.08	0.28	6460
Any Stocks	0.46	0.50	5450

Notes: This table reports means and standard deviations for additional variables used to investigate mechanisms underlying the estimated gene-wealth gradient. Each panel corresponds to an alternative mechanism. Mechanisms include Transfers (Panel A); Mortality (Panel B); Risk aversion (Panel C); Portfolio choices (Panel D); Beliefs and planning horizons (Panel E); and Defined-benefit pensions (Panel F). *This table continues onto the following page.*

Appendix Table S6: Summary Statistics: Mechanisms (continued)

	Mean	SD	N
Panel E: Beliefs and Planning Horizons	[1]	[2]	[3]
Prob: Stock Market Up			
Reported Probability	48.21	26.12	35842
Deviation from Objective	28.31	20.00	35842
Report 0%	0.05	0.22	35842
Report 50%	0.30	0.46	35842
Report 100%	0.04	0.19	35842
Prob: Major Depression			
Reported Probability	44.60	28.71	35912
Deviation from Objective	24.94	16.61	35912
Report 0%	0.07	0.26	35912
Report 50%	0.26	0.44	35912
Report 100%	0.06	0.24	35912
Prob: Double Digit Inflation			
Reported Probability	46.77	26.75	22604
Deviation from Objective	26.10	18.71	22604
Report 0%	0.06	0.23	22604
Report 50%	0.34	0.47	22604
Report 100%	0.07	0.26	22604
Planning Horizon:			
Less than 1 Year	0.13	0.34	27752
More than 1 Year	0.12	0.33	27752
More than a Few Years	0.30	0.46	27752
5-10 Years	0.34	0.47	27752
More than 10 Years	0.11	0.32	27752
Panel F: Pensions	[1]	[2]	[3]
Has DB Pension	0.57	0.49	5621
Pension Value (in $1000)	234.02	236.57	3226

Notes: This table reports means and standard deviations for additional variables used to investigate mechanisms underlying the estimated gene-wealth gradient. Each panel corresponds to an alternative mechanism. Mechanisms include Transfers (Panel A); Mortality (Panel B); Risk aversion (Panel C); Portfolio choices (Panel D); Beliefs and planning horizons (Panel E); and Defined-benefit pensions (Panel F).

B Additional Results

This appendix contains additional results referenced in the main text. Additional results are related to assortative mating, the relationship between the EA score and parental education and inheritances, and the relationship between beliefs, the financial planning horizon, wealth, and stock market participation.

B.1 Assortative Mating

Table S7 provides the evidence for the extent of assortative mating in our sample. We restrict attention to households in which both spouses have EA scores. First, we sort both males and females into quartiles, respectively, based on their individual EA scores. We then calculate the fraction of men in each male EA quartile within a given female EA quartile, and normalize so that the columns sum to one.[4] Panel A reports these distributions for individuals sorted by the raw, unadjusted value of their individual EA scores; Panel B reports distributions where individuals have been sorted based on the residual in a regression of their individual EA score on degree dummies and years of schooling. With perfect assortative mating, the matrices reported in Table S7 would be diagonal matrices, with 100% populating the diagonal entries and 0% populating the off-diagonal entries. Alternatively, random assignment would generate matrices with 25% for each entry.

We find some evidence of assortative mating, especially among the highest and lowest EA score quartiles. We find that 27.3% of women in the lowest individual EA score quartile are coupled with men in the lowest individual EA score quartile, compared to 26.9% coupled with men in the second quartile, 28.5% coupled with men in the third quartile, and 17.3% coupled with men in the fourth quartile. Entries in the fourth quartile of females' EA scores show similar patterns, with only 18.8% of the highest EA quartile women coupled with the highest EA-quartile men, compared to 33.5% coupled with the highest EA quartile males. Although we are able to reject the random-assignment null hypothesis that all entries are equal to 25% ($p < 0.001$), the degree of assortative mating appears modest relative to the counterfactual of perfect sorting. Indeed, while the within-couple correlation of years of schooling is 0.52, the within-couple correlation of individual EA scores is only 0.14.

[4]This exercise closely follows Charles, Hurst, and Killewald (2013) (see their Table 5 on p. 61), who examine assortative mating on parents' wealth.

B.2 Transfers and Parental Education

In Table S8, we relate the EA score to inheritances. All regressions include our standard controls and full education controls, unless otherwise noted. In Column [1] of Table S8, we estimate a cross-sectional regression where the dependent variable is an indicator variable equal to one if the household has ever received an inheritance over the span of the sample. In Column [2] we estimate a cross-sectional regression where the dependent variable is the log of the real dollar value of all inheritances received over the sample. Because the log of total inheritances is defined only for values greater than zero, this specification is equivalent to a regression of inheritance values conditional on receiving an inheritance. We find no relationship between the EA score and either the probability of receiving an inheritance or the size of inheritance wealth conditional on receiving an inheritance.

Next, we regress different measures of parental education on the household's average EA score. In Column [3], the dependent variable is the average education of the fathers of both household members, and we include our standard controls but no measures of respondent education. Column [4] presents the same specification but with the average of mothers' education as the dependent variable. Results suggest the education of both parents are strongly related to the household average EA score. In Columns [5] and [6], we investigate whether the relationship between parental education and the EA score is entirely explained by household members' own education. In Column [5] the dependent variable is again average fathers' education, but we now include the full set of household education controls. Column [6] reports analogous coefficients with average mothers' education as the dependent variable. The estimated coefficients on the EA score are reduced dramatically but remain statistically significant for fathers' education (and marginally significant for mothers' education), which indicates that household environments and other investments could play a role in wealth accumulation beyond just educational attainment.

B.3 Macroeconomic Beliefs and Household Behavior

The results in Table 10 suggest that individuals with lower genetic scores are more likely to report beliefs that are at odds with objective probabilities and are more likely to report "extreme" beliefs. It is possible, however, that these reported beliefs are not related to individual behavior in a meaningful way, making these results interesting but not particularly useful for understanding the potential underlying mechanisms linking the EA score to financial decisions. This would be the case if either the HRS expectations questions do a poor job of eliciting true beliefs about these economic events, or if the events themselves were not relevant for the household's choice problem. Whether a longer financial planning horizon is

associated with greater wealth is a similarly empirical question.

Table S9 shows that some of these elicited measures do indeed predict relevant behaviors such as stock market participation, and are associated with wealth. Column [1] regresses log wealth on indicators for whether the household ever reported an extreme belief for any of the three macroeconomic expectations (stock market appreciation, severe recession, or double-digit inflation), and indicators for which financial planning horizon households report. Both the financial planning horizon and stock market beliefs are related to accumulated wealth. Interestingly, excessive optimism about the stock market is actually associated with *greater* wealth, likely due to an increase in participation. This suggests that the direction of incorrect beliefs is important for their overall impact on wealth. Column [2] repeats this exercise but includes the EA score as an additional control. The inclusion of the belief and planning horizon variables reduces the coefficient on the EA score from 0.047 to 0.038.

In Columns [3] and [4] we repeat the specifications in Columns [1] and [2], but replace the log wealth with stock market participation as the dependent variable. Consistent with economic theory, longer planning horizons are associated with greater stock market participation. Reassuringly, extreme optimistic beliefs are also positively associated with stock market participation, whereas extreme pessimistic beliefs are negatively (but statistically insignificantly) related. Column [4] also shows that the positive relationship between the EA score and stock ownership documented in Table 9 remains after inclusion of belief and planning horizon controls.

Appendix Table S7: EA Score and Assortative Mating

	Panel A: Unadjusted means (N=939)				**Panel B:** Adjusted for degree and years of education (N=939)			
	Female EA Quartile				Female EA Quartile			
Male EA Quartile	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
Q1	27.3	27.4	20.3	18.8	25.4	21.6	25.4	21.9
Q2	26.9	21.4	26.3	25.9	29.0	23.4	27.2	25.3
Q3	28.5	27.4	22.8	21.9	24.6	28.8	24.6	23.6
Q4	17.3	23.9	30.6	33.5	21.0	26.1	22.8	29.1

Notes: This table reports the distribution of the male household member's EA score conditional on the quartile of the female household member's EA score for all coupled households with non-missing EA scores for both members. For each panel, each row-column entry reports the probability that a female with an EA score in the quartile corresponding to the column is coupled with a male whose EA score is in the quartile corresponding to the row. The column probabilities sum to 100 percent. Panel A presents these statistics based on unconditional individual EA scores. Panel B presents the same statistics based on the residual EA score obtained from a regression of the individual EA score on years of education and indicators for highest degree attained.

Appendix Table S8: Transfers: Inheritances and Parental Education

Dep. Var:	**Receive Inheritance**	**Inheritance Amount**	**Fathers' Educ.**	**Mothers' Educ.**	**Fathers' Educ.**	**Mothers' Educ.**
	[1]	[2]	[3]	[4]	[5]	[6]
EA Score	0.011	0.097	0.777***	0.548***	0.277***	0.111*
	(0.012)	(0.071)	(0.069)	(0.062)	(0.076)	(0.065)
Obs.	2556	1054	2294	2345	2294	2345
R^2	0.260	0.411	0.200	0.178	0.408	0.403
Standard Controls	X	X	X	X	X	X
Principal Comp.	X	X	X	X	X	X
Full Educ. Controls	X	X			X	X

Notes: This table presents estimates from regressions of inheritance and parental education variables on the average household EA score and various controls. In Column [1], the dependent variable is a binary that takes a value of 1 if the household ever receives an inheritance in the sample. In Column [2], the dependent variable is the log of the total real inheritance amount that the household receives over the course of the sample, conditional on having received an inheritance. In Column [3], the dependent variable is average years of fathers' education (averaging over household members). Column [4] repeats this exercise for average years of mothers' education. Columns [5] and [6] repeat the analysis in Columns [3] and [4], but now include controls for education of household members. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.

Appendix Table S9: Beliefs, Stock Market Participation and Wealth

Dep. Var:	**Log Wealth**	**Log Wealth**	**Own Stocks**	**Own Stocks**
	[1]	[2]	[3]	[4]
EA Score		0.038*		0.044***
		(0.021)		(0.011)
Ever Prob. Stock Mkt. Up: 0%	-0.020	-0.020	-0.044	-0.044
	(0.053)	(0.053)	(0.028)	(0.028)
Ever Prob. Stock Mkt. Up: 100%	0.169***	0.169***	0.069***	0.068***
Ever Prob. Recession: 0%	-0.000	0.000	-0.003	-0.003
	(0.037)	(0.037)	(0.021)	(0.021)
Ever Prob. Recession: 100%	0.123*	0.127*	0.011	0.015
Ever Prob. DD Inflation: 0%	-0.029	-0.026	-0.015	-0.011
	(0.045)	(0.045)	(0.025)	(0.025)
Ever Prob. DD Inflation: 100%	-0.095	-0.087	-0.017	-0.007
Min PH More than 1 Year	0.234***	0.234***	0.068***	0.066***
	(0.043)	(0.043)	(0.024)	(0.023)
Min PH More than a Few Years	0.277***	0.273***	0.078***	0.071***
	(0.040)	(0.040)	(0.026)	(0.026)
Min PH 5-10 Years	0.589***	0.588***	0.183***	0.177***
	(0.085)	(0.085)	(0.045)	(0.045)
Min PH More than 10 Years	0.761***	0.771***	0.373***	0.371***
	(0.273)	(0.278)	(0.117)	(0.120)
Obs.	5158	5158	5285	5285
R^2	0.506	0.506	0.367	0.372
Standard Controls	X	X	X	X
Principal Comp.	X	X	X	X
Full Educ. Controls	X	X	X	X
Log Income	X	X	X	X

Notes: This table presents estimates from regressions of log household wealth on measures of household subjective beliefs, planning horizons, the average household EA score, and various controls. The belief and planning horizon measures are time-invariant variables constructed from the panel of household responses. Specifically, for each of the three macroeconomic events examined in Section 5.5, we construct separate dummy variables indicating whether any household member ever reports a subjective probability of 0 percent or 100 percent, respectively. For each event, we also include the maximum deviation ever observed between a household member's subjective belief and our benchmark objective probability. We also include a series of dummy variables indicating the minimum financial planning horizon held by any household member. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.

C Robustness

This appendix contains robustness tests that are largely based on the main associations reported in Table 4. We examine robustness to the application of sampling weights, as well as alternate sample definitions, measurements of household EA scores, income specifications, wealth definitions, and control sets.

C.1 Household Structure and EA Score

This section assesses the robustness of the main results presented in Table 4 to changes in the sample definition, as well as changes in how we aggregate the EA score within two-person households. In Table S10, we address four possible selection issues that may affect the main results. In Columns [1] and [2], we repeat specifications from Columns [5] and [7] in Table 4, but apply the HRS supplied sample weights. In Columns [3] and [4] we address potential selection bias from mortality including only one observation per household; this may be important if wealthier people with higher EA scores live longer and are therefore disproportionately represented in a full panel. In Columns [5] and [6], we restrict the sample to only "coupled" households, in which the household has two members in at least one sample wave. We note that this sample may include household-year observations after a spouse has died. The results in Columns [1]-[6] all continue to demonstrate a economically large and statistically significant association between the EA score and wealth. Finally, in Columns [7] and [8], we evaluate whether the maximum or minimum EA score is ultimately driving the main association. In these specifications, we restrict the sample to households with non-missing EA scores for both members. Estimates in these specifications are quite imprecise, and in both regressions we fail to reject the null hypothesis that the coefficients on both scores are the same. The lack of precision here is likely driven by the fact that the sample size is reduced dramatically when restricting to coupled households with two non-missing scores.

C.2 Alternate Definitions of Income

The income data described in Section 3.2 are based on data from the Social Security Administration (SSA). The SSA data contain earnings information for most or all of respondents' working lives. This offers a clear advantage relative to the self-reported income measures in the HRS, which only cover older ages. However, an important limitation of the SSA data is

that they are top-coded at the taxable maximum amount for Social Security payroll taxes.[5] Panel B of Table S3 presents summary statistics relevant for the SSA income measure and top-coding. Across households, the average number of person-years with top-coded income observations is 12.67 in our sample. Less than one-third (27%) of households are never top-coded. To partially correct for top-coding, we use Current Population Survey (CPS) data to calculate mean income for people earning at least the top-coded level in each year over the period 1961-2010. We then replace the top coded amount in the SSA data with the conditional mean from the CPS data for each of these years.[6]

Table S11 presents estimates from specifications with a variety of alternative controls for lifetime household income. All specifications include the same set of controls as Column [7] in Table 4. Column [1] of S11 measures household income using the log of average household income observed within the HRS sample. Column [2] includes the log of total household income from the SSA (our standard measure), but also includes a complete set of dummy variables for each possible number of top-coded years in the SSA earnings data. Column [3] is the same as Column [2], but further includes dummy variables for each quintile of the distribution of SSA earnings across households. Column [4] controls for lifetime income using a quintic polynomial of the log of total household income from the SSA. Column [5] is the same as our basic specification (Column [7] in Table 4), but drops households in which any member is ever observed in two separate households (e.g., households that split due to a divorce). This specification addresses concerns that lifetime income might be divided across households that separate. Finally, Column [6] controls for income using the log of the average of SSA earnings over the household's 35 highest earning years. Across all of these specifications, we robustly estimate an economically large statistically significant coefficient on the EA score. Estimates of the coefficient on the EA score fall in a fairly narrow range of 0.041-0.061.

C.3 Alternative Definitions of Wealth

In this section, we repeat the analysis in Table 4 using different measures of household wealth. In Column [1] of Table S12, we use the measure of wealth provided by RAND, which does not include the present discounted value of retirement income or the retirement account balances

[5]This taxable maximum has changed substantially over time. In some years, especially in the 1960s and 1970s, a substantial portion of households fall into this category since the maximum was fairly low. For example, in 1965 the maximum was $4,800 (which is about $38,000 in 2018 dollars).

[6]For example, if an individual earned $10,000 (nominal) in 1965, we would observe a top-coded income amount of $4,800 in the SSA file. The mean CPS income for those earning at least $4,800 in 1965 is $8,103 so we would replace this individuals' income (any 1965 SSA amount of at least $4,800) with $8,103, which is approximately $56,096 in 2010 dollars.

still held with employers. In Column [2], we use the measure of wealth used in our main analysis but subtract the net value of housing. In Column [3], we again use our main wealth measure, but subtract the present discounted value of defined-benefit pension, annuity, and social security income. In Column [4], we subtract both housing and retirement-income wealth. Finally, in Column [5], we subtract the value of privately held businesses. In all specifications the key patterns from our main results remain largely unchanged.

C.4 Sample Selection

Table S13 presents estimates of our basic results from Table 4 using four different possible sample definitions. For each alternate sample, we present two specifications corresponding to Columns [5] and [7] in Table 4. Columns [1]-[2] of S13 present estimates from our baseline sample of retired households with members aged 65-75, which are reproduced from Table 4. To understand whether these results are affected by the retirement or age restrictions on this sample, we consider three other sample definitions. Columns [3]-[4] present estimates from a sample of retired households with members aged 55-85. Columns [5]-[6] present estimates from a sample of all households regardless of retirement status with members aged 50-75, while the sample used in Columns [7]-[8] includes all households with members aged ≤ 85. Across all of these samples, we consistently estimate economically large and statistically significant coefficients on the EA score in the range 0.047-0.084.

C.5 Alternative Scores

In Table S14, we examine whether the main association established in Table 4 is affected by using alternate polygenic scores for educational attainment. For each alternative score, we construct a household average and replicate Column [7] of Table 4. The score used in Column [1] is based on the Lee et al. (2018) GWAS (as is the main score used in our analysis). Since Lee et al. (2018) is the third in a series of GWAS on educational attainment by the same consortium, it is referred to as the EA3 score. The version of the EA3 score featured in Column [1] is not constructed with the LDpred method used to construct our primary score. Rather, this version is simply the sum of all SNPs weighted by their GWAS coefficients, and was released by the HRS for all genotyped waves. Consequently, this score is available for more households than the LDpred EA3 score, which was only calculated for individuals genotyped in the 2006 and 2008 waves. In Column [1] we use this non-LDpred score, but restrict the sample to households included in our main sample. We estimate a coefficient on the EA3 score of 0.038 in this specification. In Column [2], we use this score for all genotyped households (including individuals genotyped in the 2010 and 2012 waves),

and estimate a coefficient of 0.044. The stability of the coefficient across these specifications suggests that our main association is not greatly affected by the expansion of the genotyped subsample over time.

Columns [3]-[4] estimate specifications based on a score built from the Okbay et al. (2016) GWAS results (referred to as EA2, since it was the second education GWAS by this consortium), which featured a discovery sample size of $N = 293,723$. The score used in these specifications is the all-SNP (non-LDpred) score released by the HRS. In column [3] we restrict the sample to households that are included in our main sample, while column [4] expands the sample to include household-year observations with individuals genotyped in the 2010 and 2012 waves. The coefficients estimated in these specifications are similar to the coefficients estimated in Columns [1]-[2] using the EA3 scores. Finally, Columns [5]-[6] present results with the score based on the EA1 GWAS results from Rietveld et al. (2013), which featured a sample size of $N = 126,559$. Column [5] presents results for an LDpred score, while column [6] presents results from a score that sums all SNPs weighted by their GWAS association sizes. Both scores exhibit a weak, statistically insignificant association with log households wealth.

C.6 Alternative Control Sets

In this section, we examine the robustness of the main results in Table 4 to the inclusion of additional controls. In Columns [1]-[2] of Panel A in Table S15, we add the average household cognitive test score to our two baseline specifications (Columns [5] and [7] in Table 4). If the gene-wealth gradient in part arises from facility with complex decisions, a cognitive test score may explain much of the association captured by the EA score. However, the cognitive test score in the HRS is designed to capture cognitive decline and is only moderately correlated with the EA score. The average household test score is 23.89 (out of a total of 35) with averages for females and males of 24.40 and 23.22, respectively. Inclusion of the average household cognitive test score does not affect our main results. In Columns [3]-[4] of Panel A in Table S15, we include the maximum number of children associated with a household member. Higher EA score individuals may have more wealth at retirement due to having fewer children. The average number of children in the full analytical sample is 3.70, and households with higher average EA scores have fewer children. For individuals with EA scores in the first quartile, the average number of children in their household is 3.90 (again using the maximum observed for the household). For individuals with EA scores in the fourth quartile, the average is 3.34. However, inclusion of number of children leaves results unchanged.

In Panel B of Table S15, we include controls for years since retirement and years since the death of a household member. Columns [1]-[2] modify our baseline specifications by adding separate sets of dummy variables for the number of years that the male and female household members have been retired, respectively. We interact these variables with dummies for male-only and female-only households, respectively. If higher EA score individuals retire later than respondents with lower scores, this could explain greater wealth accumulation. The coefficient in column [2] decreases some, and becomes marginally statistically insignificant at the 10% level (p-value of 0.108). The death of a household member may be associated with a spike in expenses related to end-of-life care, followed by a systematic change in household consumption and decision-making. Therefore, in Columns [3]-[4], we control for the death of a household member by adding a full set of dummies for the number of years since the male and female household member has died, respectively. We also add an indicator for a coupled household that only has one member in it during a given year. The resulting coefficient estimates are again similar to our baseline estimates. Finally, in Panel C, we estimate specifications in which we simultaneously include all of the additional controls from Panels A and B. Adding all of these controls reduces the coefficient on the EA score, and in column [2] again becomes statistically insignificant.

Appendix Table S10: EA Score and Household Wealth: Robustness to Alternative Definitions of Households and EA Score Aggregation

Dep Var:								
Log Wealth	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
EA Score	0.069***	0.053**	0.081***	0.065**	0.103***	0.076***		
	(0.023)	(0.021)	(0.027)	(0.027)	(0.024)	(0.023)		
Max HH EA Score							0.052	0.034
							(0.034)	(0.034)
Min HH EA Score							0.052	0.051
							(0.036)	(0.035)
Log Income		0.265***		0.259***		0.228***		0.209***
		(0.038)		(0.046)		(0.034)		(0.056)
Obs.	5598	5286	2556	2371	3930	3723	1927	1870
R^2	0.429	0.473	0.454	0.496	0.409	0.450	0.474	0.506
Sampling Weights	X	X						
First Year Only			X	X				
Coupled HH Only					X	X		
Standard Controls	X	X	X	X	X	X	X	X
Principal Comp.	X	X	X	X	X	X	X	X
Years of Educ.	X	X	X	X	X	X	X	X
Full Educ. Controls	X	X	X	X	X	X	X	X

Notes: This table shows regression coefficients where the outcome variable is log household wealth with various control sets and sample restrictions indicated above. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.

Appendix Table S11: EA Score and Household Wealth: Robustness to Alternative Income Controls

Dep. Var: Log Wealth	[1]	[2]	[3]	[4]	[5]	[6]
EA Score	0.061***	0.041*	0.044**	0.047**	0.047**	0.051**
	(0.023)	(0.022)	(0.022)	(0.021)	(0.023)	(0.022)
Log Avg. Annual Income (HRS)	0.267***					
	(0.030)					
Log Total Income (SSA)		0.194***	0.234***	26.898	0.277***	
		(0.055)	(0.087)	(68.995)	(0.041)	
Log Total Income (SSA)^2				-4.246		
				(12.639)		
Log Total Income (SSA)^3				0.340		
				(1.120)		
Log Total Income (SSA)^4				-0.014		
				(0.048)		
Log Total Income (SSA)^5				0.000		
				(0.001)		
Log Avg. Income (Top 35)						0.260***
						(0.037)
Obs.	3993	5308	5290	5308	4895	5383
R^2	0.490	0.495	0.496	0.500	0.490	0.475
Standard Controls	X	X	X	X	X	X
Principal Comp.	X	X	X	X	X	X
Full Educ. Controls	X	X	X	X	X	X

Notes: This table shows regression coefficients where the outcome variable is log household wealth, which is regressed on average household EA score. Each column represents an alternate version of the specification in Column [7] of Table 4 with an alternate measure of household income. In Column [1], we replace our main income measure with the log of average household income for non-retired years observed in the HRS. In Column [2], we control for the number of top-coded years observed in the household's income history (separate dummy variables for each possible number of top-coded years) in addition to our main SSA measure of lifetime household income. In Column [3], we add separate dummy variables for each quintile of household income to the specification in Column [2]. In Column [4], we control for income using a quintic in the SSA log income measure. In Column [5], we use our standard SSA income measure, but restrict to households where members are never observed in multiple households. In Column [6], we control for the log of the average of SSA income in the household's highest 35 35 years of earnings. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.

Appendix Table S12: EA Score and Household Wealth: Robustness to Alternative Wealth Measures

Dep. Var: Log Wealth	RAND Wealth	Subtract Housing Wealth	Subtract Pension Wealth	Subtract Housing and Pension Wealth	Subtract Business Wealth
EA Score	0.107**	0.040*	0.126***	0.180***	0.046**
	(0.043)	(0.021)	(0.047)	(0.056)	(0.021)
Log Income	0.260***	0.285***	0.339***	0.474***	0.268***
	(0.053)	(0.037)	(0.064)	(0.069)	(0.037)
Obs.	6124	5289	4990	4657	5308
R^2	0.414	0.476	0.407	0.416	0.482
Standard Controls	X	X	X	X	X
Principal Comp.	X	X	X	X	X
Years of Educ.	X	X	X	X	X
Full Educ. Controls	X	X	X	X	X

Notes: This table shows regression coefficients where the outcome variables are different measures of log household wealth regressed onto average household EA score and various controls. Each column corresponds to the specification in Column [7] of Table 4 from the main text. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.

Appendix Table S13: EA Score and Household Wealth: Robustness to Sample Selection

Dep. Var:								
Log Wealth	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
EA Score	0.070***	0.047**	0.076***	0.060***	0.079***	0.057***	0.084***	0.067***
	(0.023)	(0.022)	(0.017)	(0.017)	(0.018)	(0.018)	(0.015)	(0.015)
Log Income		0.263***		0.217***		0.284***		0.233***
		(0.038)		(0.027)		(0.029)		(0.023)
Obs.	5621	5308	13708	12628	18925	17563	25815	23720
R^2	0.435	0.479	0.413	0.446	0.358	0.387	0.362	0.389
Standard Controls	X	X	X	X	X	X	X	X
Principal Comp.	X	X	X	X	X	X	X	X
Full Educ. Controls	X	X	X	X	X	X	X	X

Notes: This table shows regression coefficients in specifications that correspond to Columns [5] and [7] in Table 4, for different sample definitions. Columns [1]-[2] include our baseline sample of retired households with members aged 65-75, which are reproduced from Table 4. Columns [3]-[4] present estimates from a sample of retired households with members aged 55-85. Columns [5]-[6] present estimates from a sample of all households regardless of retirement status with members aged 50-75. Columns [7]-[8] include all households with members aged ≤ 85. Standard errors are clustered at the family level.

References

Charles, Kerwin Kofi, Erik Hurst, and Alexandra Killewald. 2013. "Marital Sorting and Parental Wealth." *Demography* 50 (1):51–70.

Lee, James J, Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, M Johannesson, PD Koellinger, P Turley, PM Visscher et al. 2018. "Gene Discovery and Polygenic Prediction from a 1.1-Million-Person GWAS of Educational Attainment." *Nature Genetics* .

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

Rietveld, Cornelius A, Sarah E Medland, Jaime Derringer, Jian Yang, Tõnu Esko, Nicolas W Martin, Harm-Jan Westra, Konstantin Shakhbazov, Abdel Abdellaoui, Arpana Agrawal

et al. 2013. "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science* 340 (6139):1467–1471.

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

Appendix Table S14: EA Score and Household Wealth: Robustness to Alternative Versions of the EA Score

Dep. Var: **Log Wealth**	[1]	[2]	[3]	[4]	[5]	[6]
EA3 Score no LDpred	0.038*	0.044**				
	(0.022)	(0.021)				
EA2 Score no LDpred			0.037*	0.040**		
			(0.021)	(0.020)		
EA1 LDpred					0.007	
					(0.020)	
EA1 no LDpred						0.008
						(0.020)
Obs.	5297	5964	5297	5964	5308	5308
R^2	0.479	0.474	0.479	0.474	0.478	0.478
Standard Controls	X	X	X	X	X	X
Principal Comp.	X	X	X	X	X	X
Full Educ. Controls	X	X	X	X	X	X
Log Income	X	X	X	X	X	X

Notes: This table shows regression coefficients where the outcome variable is log household wealth. In each column, we replicate Column [7] of Table 4 from the main text, but use a different version of the EA score. In Column [1], the polygenic score is constructed using all SNPs without the LDpred method based on the GWAS results of Lee et al. (2018), and the sample is restricted to include only households in our main analytical sample. Column [2] is the same as Column [1], but now expands the sample to include individuals genotyped in the 2010 and 2012 waves. In Columns[3]-[4], we use the all SNPs score based on results from a GWAS of $N = 293,723$ individuals reported in Okbay et al. (2016), which is publicly available from the Health and Retirement Study. Column [3] restricts the sample to individuals with non-missing values of the main score used in this paper, while Column [4] adds more observations by including individuals genotyped in 2010 and 2012. Columns [5]-[6] report results for scores based on the GWAS of $N = 126,559$ individuals from Rietveld et al. (2013). The score in Column [5] is based on the LDpred method, while the score in Column [6] is not.

Appendix Table S15: EA Score and Household Wealth: Robustness to Alternative Control Sets

Panel A: Dep. Var: Log Wealth	[1]	[2]	[3]	[4]
EA Score	0.058***	0.039*	0.067***	0.046**
	(0.022)	(0.021)	(0.023)	(0.022)
Avg. HH Cog. Test Score	0.032***	0.024***		
	(0.004)	(0.004)		
Max No. Children in HH			-0.041***	-0.043***
			(0.010)	(0.009)
Obs.	5495	5191	5614	5305
R^2	0.454	0.495	0.442	0.486
Log Income		X		X
Panel B: Dep. Var: Log Wealth	[1]	[2]	[3]	[4]
EA Score	0.055**	0.036	0.064***	0.041*
	(0.023)	(0.023)	(0.022)	(0.021)
Obs.	5295	5017	5621	5308
R^2	0.493	0.532	0.452	0.497
Log Income		X		X
Retirement Controls	X	X		
Mortality Controls			X	X
Panel C: Dep. Var: Log Wealth	[1]	[2]		
EA Score	0.042*	0.027		
	(0.022)	(0.021)		
Avg. HH Cog. Test Score	0.029***	0.021***		
	(0.004)	(0.004)		
Max No. Children in HH	-0.037***	-0.038***		
	(0.008)	(0.008)		
Obs.	5189	4920		
R^2	0.526	0.562		
Log Income		X		
Retirement Controls	X	X		
Mortality Controls	X	X		

Notes: This table shows regression coefficients where the outcome variable is log household wealth. Pairs of Columns for each control set correspond to those in Columns [5] and [7] of Table 4 from the main text. In Panel A, we include controls for the average cognition score in the household and the maximum number of children born to a household member. In Panel B, we add controls for years since retirement and death of a household member. In Panel C, we add all of the controls used in Panels A and B. Significance stars ***, **, and * indicate statistical significance at the 0.01, 0.05, and 0.10 levels, respectively. Standard errors are clustered at the family level.