Genes, Cognitive Skills, and Preferences

Daniel Graeber, Tanaporn Maneein, and Carsten Schröder

February 27, 2026

Abstract. We examine the relationship between cognitive genetic endowments, measured via a validated polygenic index, and economic preferences. Using representative panel data linking genetic information to stated and revealed preferences, we find that higher cognitive PGI scores are associated with greater risk aversion, increased prosociality, and higher likelihood of holding financial assets, but show no consistent relationship with patience. These associations are robust to controls for family background and population stratification. Evidence suggests that genetic endowments and parental resources act as substitutes in preference formation. Our findings highlight genetic contributions to preference heterogeneity and intergenerational economic inequality.

JEL Classification: D01, D80, D81, D90, I14, J24, J62

Keywords: genes, cognitive skills, risk aversion, patience, social preferences

---

Daniel Graeber: SOEP at DIW Berlin, IZA Bonn, and CEPA. **Tanaporn Maneein:** DIW Berlin and University of Potsdam. **Carsten Schröder:** SOEP at DIW Berlin and Freie Universität Berlin. **Corresponding author:** Daniel Graeber, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany. Email: dgraeber[at]diw.de. **Acknowledgments:** We thank Daniel Benjamin, Martin Biewen, Anna Bindler, Marco Caliendo, David Cesarini, Anthony Lepinteur, Katrine V. Løken, Levent Neyse, Thomas Siedler, Uwe Sunde, and participants at the Cluster Seminar on Public Finances and Living Conditions, the European Social Sciences and Genomics Conference 2025, the Annual Conference of the European Society for Population Economics, and the University of Potsdam for helpful comments. **Data availability:** This study uses data from SOEP-IS 2020 (DOI:10.5684/soep.is.2020), which can be accessed through the SOEP Research Data Center. **Ethics:** Data collection for this study received ethical approval from the Vrije Universiteit Amsterdam, School of Business and Economics (Application No. 20181018.1.pkr730), and from the Max Planck Society (Application No. 2019_16). **Declaration of competing interests:** The authors declare no conflict of interest.

Introduction

Preferences, particularly those related to risk, time, and social attitudes, are foundational to economic behavior and are strong predictors of education, earnings, health, and retirement choices (Fuchs 1982; Laibson 1997; Backes-Gellner et al. 2021; Lindqvist and Vestman 2011; Heckman et al. 2006; Serra-Garcia 2022). A large empirical literature also shows that measured *cognitive skills* matter for those outcomes and are systematically correlated with preference measures (Burks et al. 2009; Dohmen et al. 2010; Benjamin et al. 2013; Elinder and Erixson 2024). Because cognitive skills reflect both investments and inherited endowments, an important open question is: to what extent are genetic endowments for cognitive skills—that is, genes, which are fixed prior to any investments—related to economic preferences? Because genetic endowments are determined at conception and remain unaffected by later investments or experiences, they provide a way to study pre-market cognitive predispositions that are not shaped by confounders and substantially limit concerns about reverse causality when examining their relationship with preferences.

Understanding this link is relevant for models of human capital and the origins of inequality. If cognitive endowments and preferences co-evolve in early stages of development, their interaction may lead to greater cumulative effects on subsequent economic outcomes. This is especially relevant intergenerationally: when parental resources and genetic endowments complement each other, advantages may compound, deepening inequality across generations.

Contribution. Exploiting recent advances in social-science genomics, we provide new evidence on the relationship between a well-validated index of genetic endowments for cognitive skills, the polygenic index for cognitive skills (PGI CS), and a broad set of economic preferences, including both stated and revealed measures.[1] Using a new, nationally representative panel that links genetic data to validated preference modules, we make four contributions. First, we examine genetic variation in cognition that is not shaped by reverse causality or post-natal investments, addressing concerns affecting studies that use contemporaneous measures of cognitive skills, such as IQ tests.

Second, we document how this variation relates to risk, time, and social preferences, as well as smoking, saving, and prosocial behavior. Third, we study whether parental

[1] Throughout, we use "PGI CS" to denote the polygenic index for cognitive performance; see Becker et al. (2021).

resources amplify or mitigate these relationships, shedding light on early-life interactions between endowments and environment. Fourth, we apply Oster-style sensitivity analyses to quantify how much unobserved selection would be needed to overturn our estimates, adding a robustness tool to the social-science genomics literature. Taken together, our results, combined with the timing of genetic endowments and extensive robustness and sensitivity analyses, are consistent with a causal interpretation in which inherited cognitive endowments contribute to differences in economic preferences.

Theoretical Background. Our study builds on two theoretical foundations: the role of cognitive processes in shaping economic decision-making, and the interaction between inherited endowments and environmental investments within the skill formation framework of complementarities and substitutes.

First, models of decision-making with limited attention or cognitive resources suggest that economic preferences may partly reflect underlying differences in cognitive skills and deliberation. In these frameworks, individuals vary in the extent to which they engage in effortful, deliberative processing versus relying on heuristics or intuitive responses (Kahneman and Frederick 2002). Higher cognitive skills are associated with more consistent intertemporal choices, more accurate evaluation of risky prospects, and greater sensitivity to fairness considerations (Fudenberg and Levine 2006; Thaler and Shefrin 1981). Conversely, limited cognitive resources may generate impatience, myopic behavior, or reliance on simplification strategies such as narrow bracketing (Read et al. 1999; Rabin and Weizsäcker 2009). Related approaches model perceived utility as noisy, with cognitive complexity increasing the variance of the signal (Burks et al. 2009). These perspectives imply that cognitive limitations hinder the evaluation of complex outcomes and bias decisions toward simpler, self-focused choices in risky, intertemporal, or social contexts. Such mechanisms provide a pathway through which inherited cognitive endowments may relate to observed economic preferences and behaviors.

Second, the relationship between endowments and preferences can be framed within models of skill formation. These models emphasize that early endowments and later investments interact, potentially as complements or substitutes (Becker and Tomes 1986; Cunha and Heckman 2007; Biroli et al. 2025). When interactions are complementary, parental resources can amplify the influence of early-life endowments, potentially contributing to persistent differences in economic behavior across generations. When they are substitutes, environmental inputs may offset initial disadvantages, creating scope for policy interventions. We empirically assess whether parental background and inherited cognitive endowments operate as complements or substitutes in shaping economic preferences.

Together, these theoretical perspectives guide our empirical analysis. The decision-making frameworks provide a rationale for why cognitive endowments may be associated with economic preferences, while the skill formation literature highlights how the strength of these associations may vary with parental resources.

Data and Empirical Design. We combine genome-wide information with validated preference modules that are also used in the Global Preferences Survey (*e.g.*, Falk et al. 2018; Sunde et al. 2022; Falk et al. 2022). We study risk aversion, patience, trust, reciprocity, altruism, and a set of economically relevant behaviors consistent with these preferences, including smoking initiation, financial asset holding, and informal lending.

We proxy inherited cognitive endowments using the PGI CS, a well-validated polygenic index for cognitive skills that is constructed from genome-wide variation and explains roughly 11% of the variance in cognitive-test scores out of sample (Becker et al. 2021). For context, a recent meta-analysis finds that maternal education explains about 6.3% of the variation in cognitive skills, paternal occupation about 4.2%, and parental finances about 1.1% (Skoblow et al. 2024). Functional annotation links the underlying genetic markers to neuronal signaling pathways (Davies et al. 2018). Importantly, since genotypes are fixed prior to any individual decision-making, the PGI captures pre-market endowments that are not shaped by later experiences, although they may correlate with parental characteristics (*genetic nurture*).

Our multivariate regressions include controls for parental background, such as parental education, occupation, and separation status, as well as population stratification. These controls help address potential confounding due to family background and unobserved genetic or environmental factors; for instance, individuals with certain traits may be more likely to grow up in advantaged environments, or systematic ancestral differences could bias estimated associations (*e.g.*, Houmark et al. 2024; von Hinke et al. 2016; Benjamin et al. 2024). Importantly, evidence from Houmark et al. (2024) suggests that parental education substantially mediates the influence of parental genotypes on children's skill development. Building on this insight, including parental education, occupation, and family-structure controls helps mitigate bias arising from unobserved genetic nurture and familial genetic background. We further assess robustness using Oster-style sensitivity analyses, random-effects panel models, alternative polygenic indices (*e.g.*, for non-cognitive traits), and adjustments for multiple hypothesis testing.

In summary, these robustness checks indicate that residual confounding would need to be implausibly strong to fully account for the observed associations.

Key Results. We find that the PGI CS predicts all major preference domains, though the associations differ in strength. For risk preferences, a one standard deviation increase in the PGI CS is associated with higher stated risk aversion (+0.05 standard deviation) and with a 3.5 percentage point lower probability of smoking. For time preferences, we do not observe significant associations with stated patience, but the revealed measure of financial asset ownership shows a positive link, with a 3.4 percentage point higher probability among those with higher PGI CS. For social preferences, higher PGI CS is associated with greater stated trust (+0.09 standard deviation) and lower negative reciprocity (–0.06 standard deviation), while for revealed measures it predicts a modest increase (about 0.05 standard deviation) in the frequency of lending belongings or money to friends. By contrast, we find no robust associations with altruism or helping behavior. Estimates of Oster deltas suggest that selection on unobservables would need to be considerably stronger than selection on observables to overturn these findings. While we interpret these results with caution, the temporal precedence of genetic endowments and our series of sensitivity analyses support a causal interpretation in which inherited cognitive endowments contribute to preference formation.

By interacting the PGI CS with paternal education, a salient proxy for parental resources, we find little evidence that parental resources and genetic endowments for cognitive skills are complements. Most interaction terms are small and statistically insignificant, including trust, patience, reciprocity, and saving. An exception is risk aversion, where higher PGI CS is more strongly associated with stated risk aversion among individuals from more educated families. In contrast, our analysis suggests that genetic endowments and parental resources act as substitutes for dimensions of patience and social preferences. This substitutability offers scope for social policy to partially offset initial disparities in these endowments, thereby reducing inequalities stemming from circumstances beyond individual control and mitigating inequality of opportunity.

Literature. Our study connects four strands of literature. First, an important line of research shows that individuals with higher cognitive skills tend to be more patient, more tolerant of risk, and occasionally less pro-social (*e.g.*, Burks et al. 2009; Dohmen et al. 2010; Benjamin et al. 2013; Chen et al. 2013). However, these skills are typically measured using contemporaneous tests, such as IQ tests, which conflate innate skills with environmental influences. In contrast, our measure of innate skills, the genetic endowments for cognitive skills, is not confounded by post-natal investments or reverse causality. Our findings underscore the importance of distinguishing innate skills from environmentally influenced proxies. While proxies for cognitive skills are positively related to risk tolerance and patience, we find the opposite for risk tolerance and no relationship with patience.[2] Second, an influential literature links preferences partly to genetic factors, with twin studies estimating 10–40% heritability (Cesarini et al. 2008; Wallace et al. 2007; Cesarini et al. 2009) and molecular studies confirming weaker but consistent signals (Benjamin et al. 2012). These studies identify average genetic influence but not the mechanisms. We show that genetic endowments for cognitive skills contribute significantly to the formation of preferences. Third, a growing literature in social-science genomics uses PGIs, mostly for education, to study inequality (*e.g.*, Papageorge and Thom 2020; Barban et al. 2021; Arold et al. 2023). We instead use a PGI constructed to specifically capture cognitive skills and relate it to experimentally validated preference measures to better understand the pathway from genetic variation to economic behavior. The key distinction is that a PGI for educational attainment, a complex social outcome, captures not only cognitive skills but also non-cognitive traits, motivation, and social influences (*e.g.*, Demange et al. 2021). Fourth, we also contribute to the literature on preference formation more broadly, including work on background risk (*e.g.*, Malmendier and Nagel 2011), fairness (*e.g.*, Roth and Wohlfart 2018), and intergenerational transmission (*e.g.*, Dohmen et al. 2012).

Our study also informs the macroeconomic literature that models heterogeneity in preferences as a key source of aggregate behavior. Work in this tradition, such as Krusell and Smith (1998), Huggett et al. (2011), and Kaplan et al. (2018), emphasizes how differences in risk aversion, patience, and income dynamics shape savings, consumption, and policy transmission. However, empirical evidence on the origins of this heterogeneity remains scarce. By linking genetic endowments for cognitive skills to both stated and revealed preferences, our study provides micro-level evidence on one potential source of persistent preference heterogeneity. In doing so, we inform macroeconomic models that assume exogenous heterogeneity in preferences, offering insights into its origins and intergenerational persistence.

Our findings also speak directly to normative debates in distributive justice concerning whether individuals should be held responsible for preferences shaped by early-life

[2] Note that we are able to replicate previous studies in our data. For additional details, please refer to Appendix Section B.1.

circumstances, including genetic endowments (Rawls 1971; Dworkin 1981a,b; Fleurbaey 2008). By isolating the role of genetic variation in preference formation, we provide empirical evidence on a form of initial brute luck (Vallentyne 2002), offering a basis for evaluating justice claims around compensation for unequal economic outcomes rooted in such endowments.

The remainder of the paper details the theory on social-science genomics (Section 1), data (Section 2), empirical strategy (Section 3), results (Section 4), and concludes (Section 5).

1. Economics & Molecular Genetic Data

Before describing our empirical approach, we briefly summarize key concepts from social-science genomics that inform the construction and application of polygenic indices in applied economic research. This section provides a high-level overview of genetic variation, its measurement through single nucleotide polymorphisms (SNPs), and the genome-wide association studies (GWAS) on which polygenic indices are based. For this, we build on up-to-date reviews of Benjamin et al. (2024) and Biroli et al. (2025). Our aim is not to elaborate on biological mechanisms, but to clarify how these statistical tools produce measures of inherited endowments for cognitive skills that can be integrated into standard empirical models in economics and used to examine variation that is immune to reverse causality and not shaped by human capital investments.

Genetics. The human genome consists of deoxyribonucleic acid (DNA) comprising approximately 3 billion pairs of nucleotide molecules, which come in four types: adenine (A), cytosine (C), guanine (G), and thymine (T). The human genome is further organized into 23 pairs of chromosomes. For each pair of chromosomes, one chromosome is inherited from the father and the other from the mother. Each chromosome contains a double-stranded DNA molecule, where an "A" on one strand is always paired with a "T" on the other, and a "C" with a "G." These pairings are referred to as base pairs. Segments of these base pairs that encode specific proteins are called genes, which regulate many of the body's biological functions. The genome contains approximately 25,000 genes, which produce proteins for various functions.

Single Nucleotide Polymorphisms. Approximately 99.9% of nucleotides are identical across humans. All of the remaining differences are called polymorphisms, the locations of which are well documented. The most common type of polymorphism is the single nucleotide polymorphism (SNP), which involves a variation at a single nucleotide locus. These nucleotide variants are called alleles. Roughly 85 million SNPs have a minor allele (*i.e.*, a less common allele) with a frequency greater than 1%. Currently, millions of SNPs are documented, and many more can be imputed with high precision by leveraging the genomic correlation structure, known as linkage disequilibrium. SNPs are measured by counting the number of minor alleles, *i.e.*, variations, at a given locus, with values of 0, 1 or 2. If an individual inherits the same allele from both parents, the SNP is homozygous (coded as 0 or 2). If the alleles differ, the SNP is heterozygous (coded as 1). Most naturally occurring phenotypes (*i.e.*, outcomes which are the direct or indirect result of the expression of genes) are polygenic, meaning that many alleles jointly contribute to their expression.

Genome-Wide Association Studies. GWAS provide the foundation for linking observed genetic variants to a phenotype of interest. They estimate statistical associations between $j = 1, \ldots, J$ genetic variants (SNPs) and the outcome y. In our data, PGIs were constructed via imputation using GWAS summary statistics, as detailed in Becker et al. (2021).

The population GWAS examines all SNPs, x_{ij}, coded as 0, 1 or 2 (corresponding to the number of minor alleles), in relation to a specific outcome of subject i, y_i, controlling for baseline characteristics, z_i, such as gender, age, and the first ten principal components[3] of the genetic data, using the following model:

$$y_i = \sum_{j=1}^{J} \beta_j x_{ij} + z_i \delta + \epsilon_i. \tag{1}$$

Currently, the population GWAS in Equation (1) cannot be identified because the number of available individuals in the reference samples typically falls short of the number of SNPs. A common solution is to perform sequential regressions for each SNP j as follows:

$$y_i = \beta_j^{GWAS} x_{ij} + z_i \delta_i + \epsilon_{ij}. \tag{2}$$

Broadly speaking, $\beta_j^{GWAS} \neq \beta_j$ since z_i is typically very sparse and due to the correlation of closely spaced SNPs.[4] Importantly, conducting inference on such a large number of regressions necessitates adjustments for multiple hypothesis testing. In empirical GWAS, the common criterion for statistical significance is $p < 5 \times 10^{-8}$ (*i.e.*, 0.05 divided by 1,000,000), reflecting the approximately one million independent SNPs in the human genome.

[3]The first ten principal components of the genetic data capture the orthogonal axes which explain most of the genetic variation in a given sample (Benjamin et al. 2024, , p. 29).

[4]See Biroli et al. (2025) for further discussion.

Polygenic Indices. The small explanatory power of individual SNPs has led researchers to aggregate SNPs into polygenic indices (PGIs). PGIs are weighted linear combinations of SNPs and approximate the best linear predictor of a phenotype in independent samples. Specifically, for the construction of the PGI, one uses the estimates of the significant β_j^{GWAS} coefficients as weights for the associated SNPs. The goal is to predict as accurately as possible the propensity for an outcome in a holdout sample that was not part of the GWAS. Formally, PGIs are constructed as:

$$PGI_i = \sum_{j=1}^{J} \hat{\beta}_j^{GWAS} x_{ij}. \tag{3}$$

These PGIs generally exhibit greater predictive power than individual SNPs. In our analysis, we utilize the PGI for cognitive skills, as provided by the repository of Becker et al. (2021), to measure inherited cognitive endowments.

2. Data

We use a novel and representative dataset that integrates individual-level genetic information with detailed measures of preferences, socio-demographic characteristics, and family background: the Socio-Economic Panel-Gene (SOEP-G), which is based on the SOEP Innovation Sample (SOEP-IS).

The SOEP-IS was launched in 2011 as an extension of the German Socio-Economic Panel (SOEP), which provides longitudinally representative data on German households on an annual basis since 1984 (Goebel et al. 2019). In addition to replicating many of the contents covered by the SOEP core study, the SOEP-IS incorporates innovative content submitted through a competitive academic review process, resulting in rich measures of labor market experiences, family dynamics, attitudes, and preferences (Richter and Schupp 2015). As such, it has gained considerable attention in the economic literature (Fischbacher et al. 2024; Cobb-Clark et al. 2023; Jäger et al. 2024; Fehr et al. 2022).[5]

In 2019, 4,283 adult participants in the SOEP-IS were invited to provide saliva samples for genetic analysis in the SOEP-G study (Koellinger et al. 2023; Fraemke et al. 2025). Among these, 2,496 provided a valid saliva sample.[6] PGIs were then imputed for a range of phenotypes using the genome-wide association studies reported in Becker et al. (2021). Consistent with standard practice in the literature, individuals of non-European ancestry are excluded in quality controls, given that the majority of GWAS discovery samples are of European descent.[7] The combination of comprehensive longitudinal data with genetic information renders the SOEP-G uniquely suited for our investigation of how inherited cognitive endowments relate to economic preferences.

Genetic Endowment for Cognitive Skills. We focus on the PGI CS as our measure of inherited variation predictive of cognitive skills, that is, the component of cognitive skills not influenced by investments after conception.[8]

The PGI CS is highly predictive of a broad spectrum of cognitive test scores, with an expected out-of-sample R^2 of approximately 10.73% (Becker et al. 2021); it ranks

[5]We use the SOEP-IS 2020 (DOI:10.5684/soep.is.2020).

[6]In this study, we rely on the mildly quality-controlled sample. A sensitivity analysis using the strictly quality-controlled sample confirms our conclusions. The details about the quality controls and the sensitivity analysis are presented in Appendix Section B.6.

[7]We discuss the implications of this in the discussion in Section 5.

[8]We employ a single-trait PGI. Although multi-trait PGIs offer greater predictive power, they are derived from joint analyses of the target phenotype and additional traits (Becker et al. 2021). Consequently, multi-trait indices may capture associations stemming from phenotypic correlations that are unrelated to cognitive skills.

third in predictive power after the PGIs for height and body mass index.[9] Additional evidence indicates that the PGI CS explains significant variation in distinct cognitive domains, with incremental R^2 values of 4.64 percentage points for verbal reasoning, 4.27 percentage points for general intelligence, and 3.24 percentage points for numerical intelligence (Genç et al. 2021). Conceptually, the PGI CS can be viewed as a linear weighted average of SNPs associated with cognitive skills. Analyses of the relevant biological pathways reveal that these SNPs are linked to neurological developmental processes such as neuron projection (the extension of a nerve cell that conducts electrical impulses), neuronal differentiation, and dendrite development (the structures that facilitate the reception and propagation of electrochemical signals) (*e.g.*, Davies et al. 2018).

Stated Preferences. The stated preference measures in the SOEP have been validated in various studies, feature prominently in the Global Preference Survey (Sunde et al. 2022; Falk et al. 2018, 2022), and are extensively utilized in economics, psychology, and related disciplines.[10]

Risk preferences are measured using the question: "How do you rate yourself personally: In general, are you someone who is willing to take risks, or do you try to avoid risks?" Responses are recorded on an 11-point Likert scale ranging from 0 ("Not at all") to 10 ("Fully prepared to take risks"), and we reverse them so that higher values indicate greater risk aversion. This measure has been shown to reliably predict lottery choices and risky behaviors (Dohmen et al. 2011), and it exhibits high test–retest reliability (Lönnqvist et al. 2015).

Time preferences (patience) are assessed via the question: "Are you generally an impatient person, or someone who always shows great patience?" Responses are provided on an 11-point Likert scale from 0 ("Very impatient") to 10 ("Very patient"). For a detailed validation of this measure, see Vischer et al. (2013).

Trust is measured by three items: "On the whole, one can trust people," "Nowadays one can't depend on anyone," and "When dealing with strangers, it is better to be cautious before trusting them." Respondents indicate their level of agreement on a

[9]The expected out-of-sample predictive power of the PGIs associated with height and body mass index are 36.20% and 17.03%, respectively.

[10]For example, the risk preference item has been applied in studies on self-employment (*e.g.*, Caliendo et al. 2009), health economics (*e.g.*, Decker and Schmitz 2016; Cobb-Clark et al. 2022), household economics (*e.g.*, Serra-Garcia 2022), and intergenerational mobility (*e.g.*, Dohmen et al. 2012). Measures of trust and reciprocity are used in studies of the relevance for social preferences on economic outcomes (Dohmen et al. 2008; Fehr 2009).

scale from 1 (“Agree completely”) to 4 (“Disagree completely”). We apply principal component analysis (PCA) to the responses to these items to obtain a composite measure of generalized trust; higher factor scores indicate greater trust (see Table A1 in Appendix A for the factor loadings of trust, positive and negative reciprocity, and altruism).

Positive reciprocity is measured by three items: “If someone does me a favor, I am prepared to return it,” “I go out of my way to help somebody who has been kind to me in the past,” and “I am ready to assume personal costs to help somebody who helped me in the past.” Negative reciprocity is captured by three analogous items: “If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost,” “If somebody puts me in a difficult position, I will do the same to him/her,” and “If somebody offends me, I will offend him/her back.” Responses to both sets of items are recorded on a 7-point scale. Summary scores for positive and negative reciprocity are derived using PCA, with higher scores indicating stronger reciprocal tendencies. These measures have been validated in the economic literature (Dohmen et al. 2008).

Altruism is measured using the item “Being there for others,” for which respondents rate its importance on a scale from 1 ("Very important") to 4 ("Not at all important"). We reverse-code the responses such that higher values indicate greater altruism.

Revealed Preferences. We complement our stated preference measures with revealed behavior.[11] For risk preferences, we use an indicator for smoking behavior, following Dohmen et al. (2011). For social preferences, we consider self-reported frequencies of lending help, belongings, or money to friends. For lending help, respondents choose from four frequency options (1 “Every week”, 2 “Every month”, 3 “Less frequently”, and 4 “Never”). Meanwhile, the responses for lending belongings and money are recorded on a 5-point Likert scale (from 1, “Very often,” to 5, “Never”). We reverse these responses so that higher values correspond to more frequent lending, and normalize them to have a mean of zero and a standard deviation of one. For patience, we use self-reported ownership of financial assets as a proxy.[12]

[11]This approach follows Elinder and Erixson (2024), who rely on charitable giving, voting, and possession of eco-friendly cars as outcomes of pro-social behaviors. Note that in observational data, choices are influenced by budgets, endowments, peer effects, and other contextual factors; hence, while revealed behavior is generally consistent with stated preferences, it may also capture situational influences.

[12]Note that the item does not distinguish between risky assets, *e.g.*, stocks, and non-risky assets. However, in Germany, relatively few individuals hold stocks. In 2019, the year when the genetic information was collected, the share of individuals aged 14 and above holding any stocks was about 15.2% (Fey and Ringsleben 2024).

Genetic Principal Components. Large-scale genotyped family samples are rare, raising concerns about gene–environment correlations due to population stratification. Population stratification refers to the fact that different segments of a population, such as ethnic, religious, or regional groups, often display distinct genetic profiles (Benjamin et al. 2024; von Hinke et al. 2016). If these profiles are also associated with systematic differences in behavior or preferences, this may confound our relationships of interest. We address this issue in two ways. First, we restrict our analysis to a relatively homogeneous sample from a single European country, rather than a multi-country or continent-wide dataset. Second, we control for the first ten principal components (PCs) of the genetic data, which capture the major axes of genetic variation in our sample (Menozzi et al. 1978). In population genetics, such PCs approximate differences that emerge when subgroups diverge over time due to random changes in allele frequencies.[13] Accordingly, we include the first ten genetic PCs in all regressions to adjust for population stratification.

Parental Background. Controlling for family background is essential to account for gene–environment associations, which arise because parents and children share genetic endowments and because parents with higher cognitive endowments may provide more favorable environments for the development of cognitive skills. We include indicators for maternal and paternal educational attainment, parental occupational status at age 15, and whether the parents lived separately until the respondent reached age 15.[14] In related applications, Houmark et al. (2024) show that conditioning on realized parental outcomes may sufficiently adjust for genetic nurture effects.[15]

Sample Construction. For the main analyses, we combine the time-invariant genetic information from 2019 with data on preferences and sociodemographics from the same year or the closest available year. Summary statistics for the key variables are reported in Table 1. Figure 1 displays the distribution of the PGI CS.[16] All variables are presented in detail in Tables A2 and A3.

[13]Genetic drift refers to random changes in allele frequencies over generations in a finite population. When groups cease interbreeding, genetic drift can generate differences in allele frequencies and in the correlational structure of SNPs, which are reflected in principal components (Benjamin et al. 2024).

[14]A separate category is created for respondents reporting "Do not know" for each of these controls.

[15]We discuss the importance of these controls in more detail in Section 3.

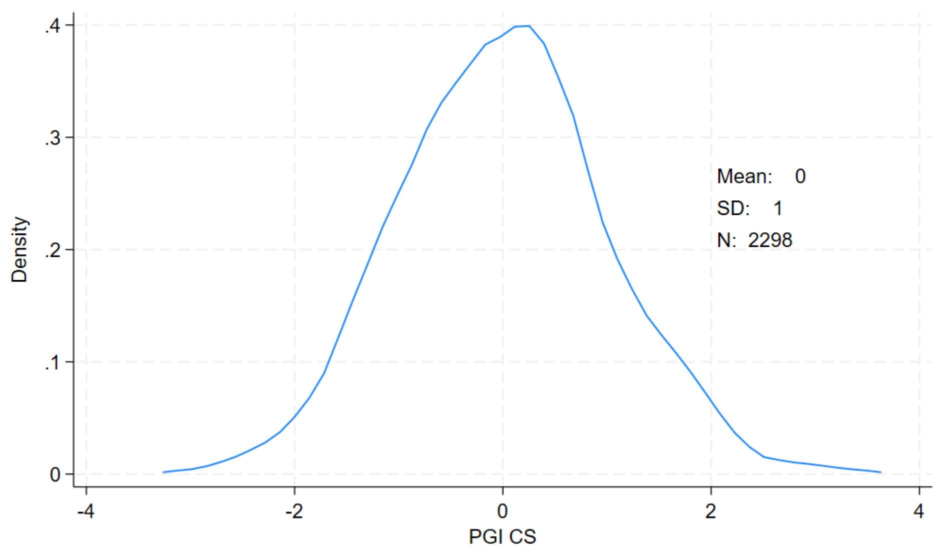
[16]While this is an ad hoc choice, merges with alternative years are possible. We assess the sensitivity of our results to these design choices in Section 4.2 and provide further details in Appendix B.3.

Table 1. Summary statistics

	Mean	SD	N		Mean	SD	N
Stated preferences				Mother's occupation:			
Trust	0.00	1.00	2,189	Self-employed	0.06	0.24	2,298
Positive reciprocity	-0.00	1.00	2,197	White-collar	0.33	0.47	2,298
Negative reciprocity	-0.00	1.00	2,197	Civil service	0.02	0.13	2,298
Altruism	2.38	0.56	2,130	Blue-collar	0.18	0.39	2,298
Patience	5.89	2.53	2,200	Not employed	0.32	0.47	2,298
Risk aversion	5.04	2.36	2,298	Missing	0.09	0.29	2,298
				Mother's education:			
Revealed preferences				Upper secondary sch.	0.08	0.28	2,298
Smoking	0.23	0.42	2,057	Intermediate sch.	0.20	0.40	2,298
Financial assets	0.45	0.50	2,256	Technical sch.	0.01	0.10	2,298
Lending help	1.64	0.83	2,173	Secondary sch.	0.58	0.49	2,298
Lending belongings	1.26	1.01	2,166	No school degree	0.03	0.18	2,298
Lending money	0.59	0.82	2,167	Other degree	0.02	0.14	2,298
				Missing	0.07	0.26	2,298
Socio demographic				Father's education:			
Geburtsjahr	1,963.97	19.20	2,298	Upper secondary sch.	0.15	0.35	2,298
Female	0.54	0.50	2,298	Intermediate sch.	0.15	0.36	2,298
Father's occupation:				Technical sch.	0.01	0.09	2,298
Self-employed	0.12	0.32	2,298	Secondary sch.	0.55	0.50	2,298
White-collar	0.29	0.46	2,298	No school degree	0.02	0.15	2,298
Civil service	0.08	0.27	2,298	Other degree	0.02	0.14	2,298
Blue-collar	0.35	0.48	2,298	Missing	0.10	0.31	2,298
Not employed	0.09	0.29	2,298	HH. env. growing up:			
Missing	0.07	0.26	2,298	Two-parents HH.	0.73	0.44	2,298
				Other types of HH.	0.22	0.42	2,298
				Missing	0.05	0.22	2,298

Notes: Table 1 displays the summary statistics for the main variables used in the analysis. The analytical sample includes 2,298 unique individuals from SOEP-IS. The indices measuring trust and reciprocity are derived from principal component analysis (PCA) and have been standardized to zero mean and unit variance. Authors' calculations based on SOEP-IS 2020.

Figure 1. Distribution of PGI for cognitive skills



Notes: The PGI for CS is standardized to have mean zero and standard deviation one in our sample. The kernel density estimate is obtained using an Epanechnikov kernel with 0.1872 bandwidth. Authors' calculations based on SOEP-IS 2020.

3. Empirical Strategy

We examine whether inherited cognitive endowments, as proxied by the PGI CS, are systematically related to economic preferences. While the analysis is based on observational data, the temporal ordering of genetic variation and the inclusion of rich controls allow for a cautious interpretation of the estimates as capturing the contribution of pre-market cognitive endowments to observed preference heterogeneity. Our empirical strategy is based on the following model:

$$y_i = \beta_0 + \beta_1 PGI\ CS_i + \boldsymbol{\beta_2 X_i} + \boldsymbol{\beta_3 Z_i} + \boldsymbol{\beta_4 PC_i} + \eta_i. \tag{4}$$

In Equation (4), y_i denotes the preference of individual i in one of three domains: risk, time, or social preferences. The key explanatory variable, $PGI\ CS_i$, is the PGI for cognitive skills. Conditional on a rich set of covariates, including demographic characteristics ($\boldsymbol{X}_i$), family background ($\boldsymbol{Z}_i$), and population structure ($\boldsymbol{PC}_i$), the coefficient β_1 is interpreted as capturing the reduced-form relationship between inherited cognitive endowments and preferences.

Specifically, $\boldsymbol{X}_i$ includes gender, a second-order polynomial in age (measured in the survey year when the outcome is observed), and year-of-birth fixed effects. $\boldsymbol{Z}_i$ accounts for family background, including parental education (measured by school-leaving degrees), occupational status for each parent, and an indicator for parental separation. $\boldsymbol{PC}_i$ includes the first ten PCs of the genetic data to adjust for population stratification. We report robust standard errors throughout.

3.1. Addressing Potential Confounders

A causal interpretation of $\hat{\beta}_1$ relies on two key assumptions commonly invoked in observational studies: *conditional independence* and a *conditional exclusion restriction*.

More precisely, conditional on the control variables included in Equation (4), we would need to assume that the PGI for cognitive skills is uncorrelated with unobserved determinants of preferences (conditional independence), and that the PGI affects preferences primarily through its association with cognitive skills rather than through other pathways (conditional exclusion restriction). Violations of these assumptions would imply that $\hat{\beta}_1$ may capture not only the contribution of inherited cognitive endowments but also confounding influences, such as environmentally mediated gene–environment correlations (*e.g.*, more favorable home environments associated with higher parental cognitive skills) or residual population stratification. Below, we discuss these assumptions in greater detail and describe the strategies we employ to mitigate potential sources of bias. Notably, while we remain cautious in interpreting our estimates as causal, our robustness analyses suggest that bias from unobserved confounders would need to be implausibly large relative to the observed association with the PGI to explain away our results (*e.g.*, Oster 2019; Altonji et al. 2005).

Conditional Independence. The conditional independence assumption requires that, conditional on a sufficiently rich set of covariates, the potential outcomes for preferences are independent of $PGI\ CS_i$. While this assumption is fundamentally untestable, because it pertains to counterfactual outcomes, violations may arise if allele frequencies vary systematically across population subgroups (von Hinke et al. 2016; Benjamin et al. 2024), and if these subgroups also differ in preferences due to correlated environmental factors (*e.g.*, cultural or religious norms). To mitigate this concern, we restrict our sample to a relatively homogeneous population consisting of individuals residing in Germany, as opposed to individuals living in Europe or the U.S., and include genetic PCs ($\boldsymbol{PC}_i$) to adjust for residual population structure. These components capture major axes of ancestral genetic variation and serve as controls for potential stratification bias.

Conditional Exclusion Restriction. For $\hat{\beta}_1$ to admit a causal interpretation, we must assume that $PGI\ CS_i$ influences preferences primarily through its relationship with cognitive skills, conditional on observable controls. This conditional exclusion restriction may be violated if individuals with higher cognitive PGIs also experience systematically different environments, for example, if parents who transmit advantageous genetic endowments also invest more in extracurricular activities or other aspects of child-rearing that directly shape preferences. The ideal strategies to address this concern would involve within-family designs or parental genotype data; however, our sample, like most with available genetic information, includes few sibling pairs and lacks comprehensive data on parental genotypes.

As an alternative to controlling directly for parental genotypes, we include a detailed set of parental background variables in $\boldsymbol{Z}_i$, which proxy for the socio-economic environment during the respondent's upbringing. Specifically, we control for parental education, parental occupational status, and an indicator for parental separation before age 15. Recent evidence by Houmark et al. (2024) suggests that conditioning on parental education substantially attenuates environmental confounding in analyses of child outcomes when parental genetic data are unavailable.[17]

To assess the importance of these controls, Table A4 in Appendix A compares background characteristics for individuals above and below the median of the PGI CS distribution. We report means (for continuous variables), shares (for categorical variables), and standardized differences. Notable imbalances across these groups underscore the importance of controlling for family background and genetic PCs. However, as von Hinke et al. (2016) notes, imbalance in observed covariates does not necessarily imply a violation of conditional independence or the exclusion restriction, as some background characteristics may lie on the causal pathway from inherited endowments to preferences.

To further probe potential gene–environment correlations, Table 2 presents regressions of two proxy measures for cognitive skills on $PGI\ CS_i$: years of schooling (Panel A) and an indicator for obtaining A-levels[18] (Panel B). Given the limited availability of test-based measures of cognitive skills in our data, we rely on educational attainment as a proxy. Each panel sequentially adds controls: starting with baseline demographics and genetic PCs (column 1), then adding parental education (column 2), occupational status (column 3), and parental separation (column 4). The full specification is shown in column 5. For years of schooling (Panel A), including parental education reduces the coefficient on $PGI\ CS_i$ by approximately 25%; adding parental occupation reduces it by a further 12%, while parental separation has little effect. Similar patterns hold for the A-level outcome (Panel B). These results highlight the importance of accounting for family background when interpreting the estimated relationship between the cognitive PGI and preferences, and help to mitigate concerns about violations of the exclusion restriction.

[17] Houmark et al. (2024) study the interaction between children's polygenic indices for educational attainment and parental investments in skill formation.

[18] A-levels correspond to the German *Abitur*, the qualification required for university entrance.

Table 2. Association between PGI CS and schooling

	(1)	(2)	(3)	(4)	(5)
Panel A: Years of schooling					
PGI CS	0.337***	0.254***	0.297***	0.335***	0.250***
	(0.036)	(0.034)	(0.034)	(0.036)	(0.034)
Parental education	No	Yes	No	No	Yes
Parental occup. status	No	No	Yes	No	Yes
Parents separated	No	No	No	Yes	Yes
R^2	0.133	0.264	0.198	0.144	0.278
Adjusted R^2	0.095	0.228	0.159	0.107	0.239
Panel B: Probability of obtaining A levels					
PGI CS	0.092***	0.070***	0.082***	0.092***	0.069***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Parental education	No	Yes	No	No	Yes
Parental occup. status	No	No	Yes	No	Yes
Parents separated	No	No	No	Yes	Yes
Observations	2227	2227	2227	2227	2227
R^2	0.119	0.233	0.172	0.130	0.245
Adjusted R^2	0.081	0.195	0.132	0.092	0.204

Notes: Table 2 displays the effect of the polygenic index for cognitive skills on years of schooling (in Panel A) and on the probability of obtaining A levels (in Panel B). Column (1) only controls for gender, age, year of birth, and the first ten principal components of the genetic data. Column (2) adds a full set of indicators for parental educational degrees. Column (3) adds indicators for parental occupational status. Column (4) adds years living with both parents until age 15. Column (5) includes all controls from models (2) to (4). The polygenic index for cognitive skills is standardized with mean zero and standard deviation one. Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

4. Results

Subsection 4.1 presents our main findings. Leveraging variation in genetic endowments for cognitive skills and conditioning on detailed family and ancestry controls, we examine the relationship between the PGI for cognitive skills and preferences. We find that the PGI for cognitive skills is positively associated with risk aversion and social trust, and negatively associated with negative reciprocity. Subsection 4.2 reports a series of complementary analyses addressing potential sources of bias, including omitted variables, multiple hypothesis testing, the quality of the genetic data, model specification, and sample selectivity. These analyses suggest that our results are robust across a broad set of statistical and measurement concerns. Finally, Subsection 4.3 investigates interactions between the PGI and parental resources. We find some evidence of complementarities for risk aversion, but little evidence of such complementarities for time or social preferences. For some outcomes, our findings are consistent with substitutability between genetic endowments and early-life environmental resources.

4.1. Main Results

This section examines the relationship between inherited cognitive endowments and economic preferences. Figures 2 and 3 present binned scatterplots illustrating the unconditional relationships between the PGI for cognitive skills and various preference measures. Our main estimates, which aim to assess this relationship under the empirical strategy outlined in Section 3, are reported in Tables 3 and 4 and summarized graphically in Figure 4.

For each preference measure, the regression results reflect estimates from four specifications of Model (4), which progressively introduce controls to address potential confounding. Panel A reports estimates controlling for baseline demographic characteristics (gender, year of birth, and age). Panel B adds controls for population structure using genetic principal components. Panel C further includes family background variables. Panel D presents our preferred, fully saturated specification. By simultaneously controlling for demographics, genetic PCs, and parental background, this model provides a robust assessment of the relationship between cognitive genetic endowments and preferences.

Risk Preferences. The upper panel of Figure 2 shows the unconditional relationships between our risk preference measures and the PGI for cognitive skills. The association between stated risk aversion and the PGI CS is positive but modest. A linear fit, depicted by the orange line, suggests that a one standard deviation increase in the PGI CS is associated with a 0.058 standard deviation increase in stated risk aversion. For smoking, the unconditional relationship is negative: a one standard deviation increase in the PGI CS is linked to a 4.8 percentage point decrease in the probability of smoking.

Table 3 reports estimates from our main empirical model, providing a more comprehensive view of the relationship between cognitive genetic endowments and risk preferences. For stated risk aversion, the estimate from the baseline specification in Panel A indicates that a one standard deviation increase in the PGI CS is associated with a 0.043 standard deviation increase in risk aversion. This estimate is stable across specifications, increasing slightly to 0.046 with the inclusion of genetic PCs (Panel B), and to 0.047 with the addition of family background controls (Panel C). In the fully saturated model (Panel D), which includes all controls, a one standard deviation increase in the PGI CS is associated with a 0.050 standard deviation increase in stated risk aversion.

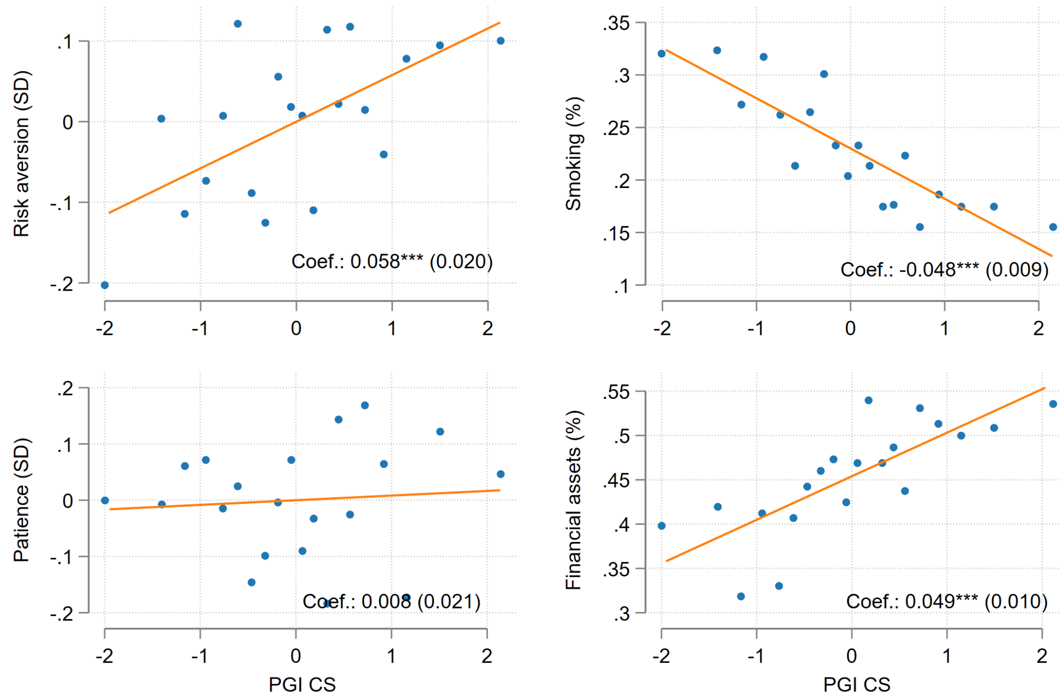
We find similar patterns for smoking behavior. The estimate from our preferred specification (Panel D) suggests that a one standard deviation increase in the PGI CS is associated with a 3.5 percentage point reduction in the probability of smoking, equivalent to a 15.22% reduction relative to the baseline mean. While the inclusion of genetic PCs does not materially change the estimate (compare Panel B to Panel A), adding family background controls slightly reduces the magnitude from –0.040 to –0.035 (Panel C).

Taken together, these results are consistent with higher inherited cognitive endowments being systematically related to greater risk aversion and a lower likelihood of smoking. The observed variation across specifications underscores the importance of accounting for potential omitted variables and gene–environment correlations. Ignoring these factors could bias the estimated associations.

Patience. We assess the relationship between cognitive endowments and two measures of patience: stated patience and the probability of holding financial assets. For stated patience, we find no evidence of a systematic association. Across all specifications, including our preferred model (Panel D of Table 3), the estimated coefficient on the PGI CS is small in magnitude and statistically insignificant, suggesting no meaningful relationship.

The findings differ for financial asset holding. The unconditional relationship is positive, with a linear coefficient of 0.049. Consistent with this, the baseline specification

Figure 2. Unconditional associations between the PGI CS and risk aversion and patience



Notes: Figure 2 displays binned scatter plots of risk aversion, the probability of smoking, patience, and the probability of holding financial assets against the PGI CS. The linear fit stems from an OLS regression of the outcomes on the PGI CS. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Authors' calculations based on SOEP-IS 2010–2020 (DOI:10.5684/soep.is.2021).

(Panel A of Table 3) estimates that a one standard deviation increase in the PGI CS is associated with a 4.7 percentage point increase in the probability of holding financial assets. This association remains stable after adding genetic PCs (Panel B), with an estimated effect of 4.3 percentage points. Adding family background controls (Panel C) leads to a modest reduction to 3.6 percentage points. The fully specified model (Panel D), which includes all controls, yields an estimated association of 3.4 percentage points.

Overall, moving from the most parsimonious to the most comprehensive specification reduces the estimated association by approximately 28%, suggesting some role for confounding via family background. Taken together, our analysis yields a nuanced picture. While the PGI for cognitive skills appears unrelated to stated patience, it is positively associated with the likelihood of holding financial assets. Although this divergence may appear inconsistent at first glance, it likely reflects the multidimensional nature of patience. The stated patience measure captures general self-assessed patience without context, whereas financial asset holding is a behavioral measure of financial patience. It is therefore plausible that individuals differ in their expression of patience

Table 3. Association between PGI CS and risk/time preferences

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
Panel A: Demographic controls				
PGI CS	0.043**	-0.040***	-0.004	0.047***
	(0.020)	(0.009)	(0.022)	(0.011)
R^2	0.125	0.114	0.046	0.073
Adjusted R^2	0.092	0.077	0.010	0.039
Panel B: Demographic and genetic controls				
PGI CS	0.046**	-0.040***	-0.002	0.043***
	(0.020)	(0.009)	(0.022)	(0.011)
R^2	0.130	0.118	0.050	0.084
Adjusted R^2	0.094	0.076	0.010	0.047
Panel C: Demographic and family controls				
PGI CS	0.047**	-0.035***	-0.001	0.036***
	(0.020)	(0.009)	(0.022)	(0.011)
R^2	0.133	0.140	0.055	0.106
Adjusted R^2	0.091	0.093	0.008	0.063
Panel D: Full set of controls				
PGI CS	0.050**	-0.035***	0.003	0.034***
	(0.021)	(0.009)	(0.023)	(0.011)
Observations	2298	2057	2200	2256
R^2	0.138	0.143	0.061	0.113
Adjusted R^2	0.093	0.092	0.009	0.067

Notes: Table 3 displays the results from regressing risk/time preference measures on the polygenic index for cognitive skills under different specifications. Panel A shows the most parsimonious model, controlling for gender, year of birth, age, and age squared. Panel B further controls for the first ten principal components of the genetic data. Panel C adds parental educational attainment, occupational status, and an indicator for parental divorce to the basic control set. Panel D combines all these controls. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

across different life domains.

Social Preferences. Our analysis suggests that genetic endowments for cognitive skills are systematically related to different facets of social preferences, including trust, reciprocity, and lending behavior. For trust, both the unconditional relationships in Figure 3 and the regression estimates in Table 4 point to a positive association with the PGI CS. In our preferred specification (Panel D), a one standard deviation increase in the PGI CS is associated with a 0.092 standard deviation increase in trust.

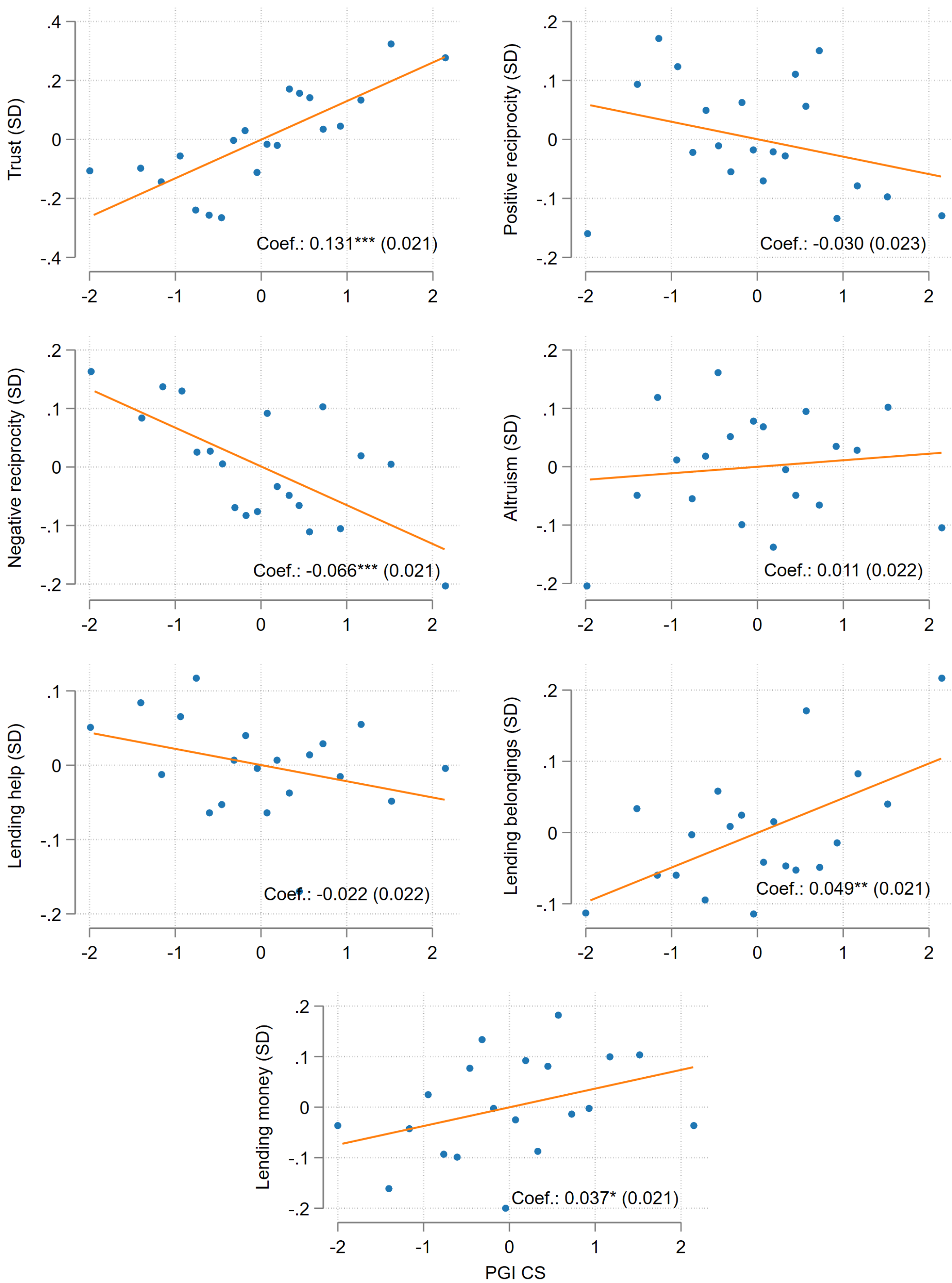
Results for reciprocity are mixed. We find no evidence that the PGI CS is related to positive reciprocity, whereas the estimates indicate a robust negative association with negative reciprocity, with coefficients ranging from –0.065 to –0.056 across Panels A through D. For altruism and lending help, all specifications consistently show no meaningful relationship with the PGI CS. By contrast, we find stable evidence that higher PGI CS is associated with increased frequency of lending belongings and lending money, suggesting that individuals with higher cognitive genetic endowments engage more frequently in pro-social material support behaviors.

TABLE 4. Association between PGI CS and social preferences

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
Panel A: Demographic controls							
PGI CS	0.135***	-0.032	-0.063***	0.023	-0.003	0.077***	0.063***
	(0.021)	(0.023)	(0.021)	(0.022)	(0.022)	(0.021)	(0.020)
R^2	0.057	0.036	0.061	0.083	0.105	0.116	0.192
Adjusted R^2	0.022	-0.001	0.024	0.048	0.071	0.082	0.162
Panel B: Demographic and genetic controls							
PGI CS	0.121***	-0.020	-0.065***	0.023	-0.005	0.069***	0.066***
	(0.022)	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)	(0.020)
R^2	0.076	0.044	0.065	0.093	0.113	0.120	0.195
Adjusted R^2	0.036	0.002	0.024	0.054	0.075	0.082	0.161
Panel C: Demographic and family controls							
PGI CS	0.104***	-0.039	-0.056**	0.017	-0.009	0.058***	0.052**
	(0.021)	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)	(0.020)
R^2	0.104	0.045	0.074	0.094	0.122	0.137	0.208
Adjusted R^2	0.060	-0.003	0.027	0.049	0.078	0.094	0.169
Panel D: Full set of controls							
PGI CS	0.092***	-0.025	-0.058***	0.016	-0.009	0.053**	0.055***
	(0.022)	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)
Observations	2189	2197	2197	2130	2173	2166	2167
R^2	0.118	0.053	0.079	0.104	0.130	0.140	0.211
Adjusted R^2	0.070	0.000	0.027	0.054	0.082	0.093	0.168

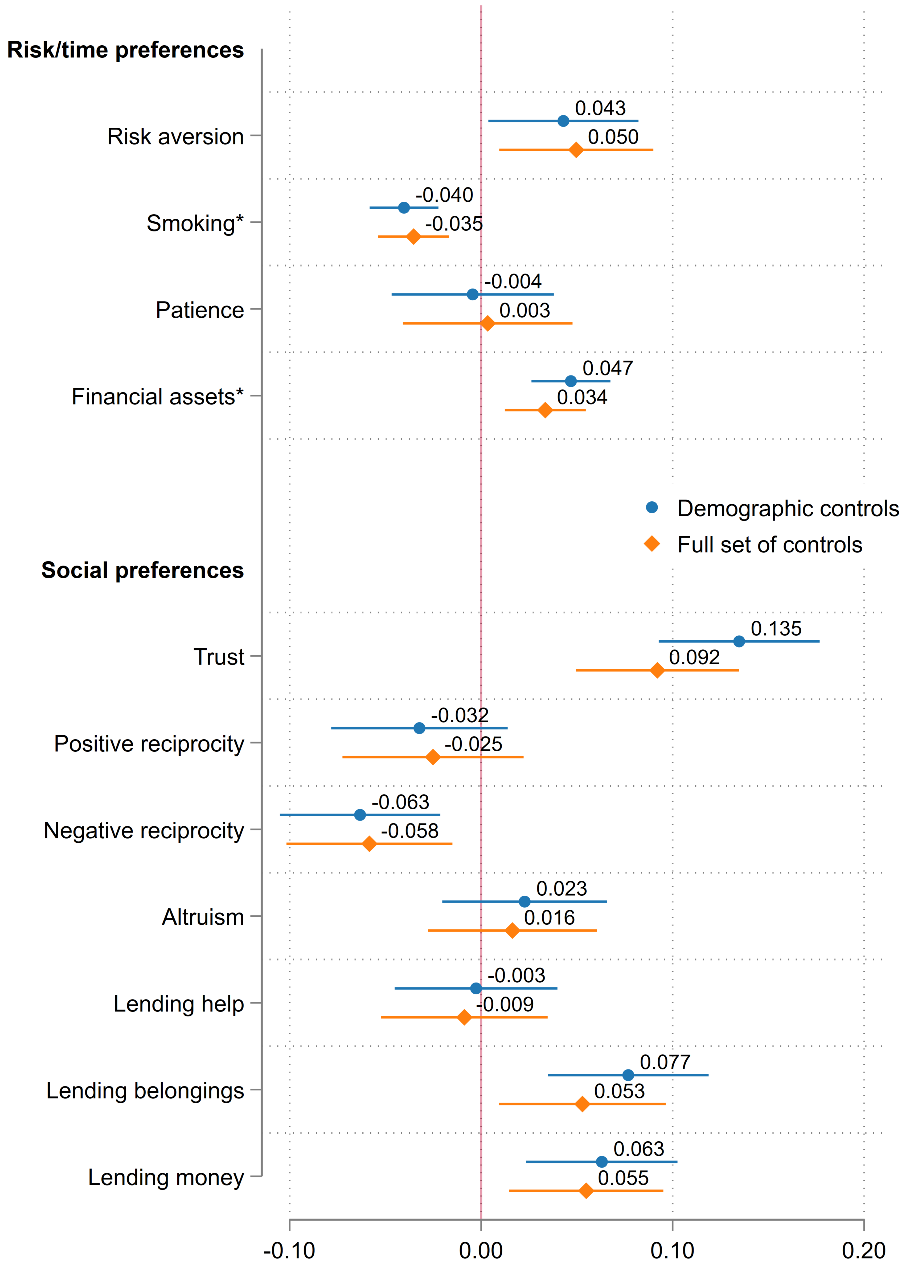
Notes: Table 4 displays the results from regressing social preference measures on the polygenic index for cognitive skills under different specifications. Panel A shows the most parsimonious model, controlling for gender, year of birth, age, and age squared. Panel B further controls for the first ten principal components of the genetic data. Panel C adds parental educational attainment, occupational status, and an indicator for parental divorce to the basic control set. Panel D combines all these controls. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

FIGURE 3. Unconditional associations between PGI CS and social preference outcomes



Notes: Figure 3 displays unconditional associations between the PGI CS and outcomes for trust, positive reciprocity, negative reciprocity, altruism, and frequencies of lending help, belongings, and money. The linear fit is based on OLS regressions of outcomes on PGI CS. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Authors' calculations based on SOEP-IS 2010–2020.

FIGURE 4. The effect of the PGI CS on social preferences



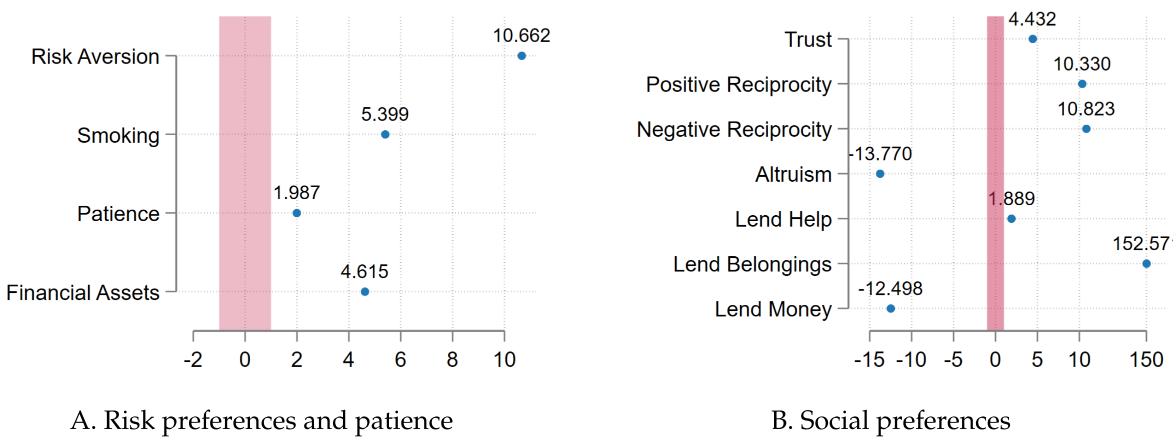
Notes: Figure 4 shows the association of the PGI CS with preferences outcomes. The estimates in blue, represented by the dots, are obtained from a regression of the respective outcome on the PGI CS, controlling for gender, year of birth, and age. The orange estimates, represented by diamonds, are derived from a regression of the respective outcome on the PGI CS, with additional controls for gender, year of birth, age, the first ten principal components of the genetic data, parental education, parental occupational status, and parental separation status. All outcome are measured in standard deviations if not otherwise noted. An asterisk beside the outcome indicates that these outcomes are on the probability scale. Horizontal bars indicate 95% confidence intervals based on robust standard errors. Authors' calculations based on SOEP-IS 2010-2020.

4.2. Robustness Analyses

Omitted Variables. Our empirical model (Equation (4)) includes an extensive set of covariates, which reduces concerns about omitted variable bias. Nevertheless, two potential sources of endogeneity are particularly relevant. First, unobserved factors may jointly influence both the PGI for cognitive skills and preference outcomes, biasing our estimates. Second, the PGI CS may be linked to preferences through indirect behavioral channels,such as genetic nurture, that also affect cognitive development, thus violating the assumption of conditional exogeneity.

To assess the potential severity of omitted variable bias, we apply the method proposed by Oster (2019), estimating the *delta* parameter that captures the relative degree of selection on unobservables (vs. observables) required to explain away the estimated effect. Following Altonji et al. (2005) and Oster (2019), values of $\delta > 1$ suggest robustness to omitted variable bias.[19] The results, shown in Tables A6 and A7 and visualized in Figure 5, indicate high robustness. For outcomes with statistically significant associations, including risk aversion, trust, negative reciprocity, smoking, financial asset holding, and lending belongings or money, the δ values range from approximately 5 to 10. This suggests that unobserved selection would need to be substantially stronger than selection on observables to account for the estimated relationships.

FIGURE 5. Robustness to omitted variable bias using Oster (2019)'s δ-method



Notes: Figures 5A and 5B display estimates of δ as proposed by Oster (2019), reflecting the strength of unobserved selection required to explain away the estimated effect. Values $\delta > 1$ indicate robustness. The x-axis in Figure 5B is truncated (10–150) for readability.

[19]Methodological details are presented in Appendix Section B.2.

Unobserved Time-Invariant Heterogeneity. To address concerns about time-invariant unobserved individual heterogeneity, such as unmeasured family background or latent population substructure, we exploit the panel dimension of our data. While fixed effects estimation is not feasible due to the time-invariant nature of the PGI CS, we estimate random effects models using all available survey waves.[20] This approach allows us to absorb unobserved individual heterogeneity and approximate a control for unmeasured genetic background, including parental genotype, which is relevant in the context of Mendelian randomization.[21,22] The results, presented in Tables A8 and A9, are consistent with our main findings, supporting the stability of the estimated relationships.

Pleiotropy. Another potential threat to interpretation is pleiotropy, where genetic variants influence multiple traits (Benjamin et al. 2024). To address this, we re-estimate our main model (Equation (4)) including polygenic indices for three non-cognitive traits available in the SOEP-G: openness to experience, extraversion, and neuroticism. Tables A10 and A11 (Appendix B.4) show that the estimated associations between the PGI CS and preferences remain largely unchanged. Thus, pleiotropic pathways via these personality traits do not appear to account for the observed patterns. Since the PGIs for cognitive and non-cognitive skills may include SNPs in linkage disequilibrium, this analysis also addresses potential correlation between genetic predictors.

Inference. Our analysis involves testing for associations across multiple outcomes (up to eleven hypotheses), which increases the risk of *false discoveries*. To address this concern, we apply multiple hypothesis testing adjustments tailored to different families of outcomes. Tables A12 and A13 in Appendix Section B.5 summarize the results. Specifically, Table A12 reports adjusted p-values based on three procedures: (a) the method proposed by List et al. (2019), (b) a Bonferroni correction (*e.g.*, Dunn 1961), and (c) the stepwise adjustment procedure of Westfall and Young (1992).[23] Table A13 replicates this analysis while grouping outcomes by preference domain. For example, risk aversion and smoking are treated as one hypothesis family, and stated patience and asset holding as another. Across most of the main outcomes, the estimated associations remain

[20]Outcome availability varies by wave; see Appendix B.3 for details.

[21]Mendelian randomization relies on random allocation of alleles during meiosis; see Benjamin et al. (2024).

[22]Random effects models assume no correlation between unobserved individual effects and regressors (Wooldridge 2010).

[23]Further details are provided in Appendix Section B.5.

statistically significant under at least two of the three correction methods, reinforcing the robustness of the results to concerns about multiple testing.

Measurement of Genetic Information. To assess whether data quality affects the validity of our findings, we replicate our main analyses using a subsample subjected to stricter quality control procedures. The baseline genetic data come from the SOEP-G sample, processed using standard quality control (QC) protocols. In addition, Koellinger et al. (2023) implement an enhanced QC procedure based on per-chromosome missingness, discrepancies between self-reported and genetically inferred sex, and deviations in heterozygosity and homozygosity.[24] Tables A14 and A15 report the results based on the stricter QC sample. The estimated associations are nearly identical to those based on the standard QC sample, suggesting that measurement quality in the genetic data does not materially affect the results.

Another relevant consideration is that the PGI for cognitive skills (PGI CS) is an empirical proxy for the underlying genetic influence on cognitive ability, as the true population-level GWAS is unobserved. As a result, the PGI CS is likely measured with *classical measurement error*, which attenuates the estimated associations toward zero (Becker et al. 2021; Benjamin et al. 2024). To assess and correct for this potential attenuation bias, we follow the procedure proposed by Becker et al. (2021) (see Appendix Section B.7 for details). This method relies on an adjustment factor analogous to a reliability ratio, which is not available for the SOEP-G sample. We therefore draw on external estimates from other datasets,specifically the Wisconsin Longitudinal Study (WLS), Add Health, and UK Biobank, to approximate the necessary correction.[25]

Figure A1, Figure A2, and Table A16 present the corrected estimates. As expected, the point estimates increase in magnitude across all outcomes while maintaining the same sign, though the associated confidence intervals widen. In many cases, the confidence intervals overlap with those from the uncorrected estimates. This pattern suggests that while attenuation bias may lead to conservative estimates in our baseline model, the differences are not statistically significant. Taken together, the results provide suggestive evidence that the true associations between cognitive endowments and preferences may be larger than our baseline estimates, potentially by a factor of two to three, but remain within the bounds of statistical uncertainty.

[24]Heterozygosity refers to the likelihood that two alleles at a locus differ within an individual (Nei 1973); homozygosity is its inverse. See Appendix Section B.6 or Koellinger et al. (2023) for details.

[25]These external studies provide empirical estimates of the predictive accuracy of the PGI for cognitive skills.

Model Specification. Two aspects of model specification warrant further consideration. First, the assumption of a linear relationship between the PGI for cognitive skills and preferences may be too restrictive. Second, some behavioral outcomes may be influenced by mediating factors such as income and education. To relax the linearity assumption, we re-estimate Equation (4) including a second-order polynomial in the PGI CS. The results, reported in Tables A17 and A18 (Appendix B.8), show that nonlinearity is not statistically significant for most outcomes. Exceptions include trust and the frequency of lending belongings, where the relationship appears convex, suggesting that the association increases at higher levels of the PGI CS.

Regarding potential mediation, our baseline models exclude income and education to avoid introducing collider bias. These variables are likely influenced by both the PGI CS and unobserved factors that may also affect preferences, making them unsuitable as standard controls (Rosenbaum 1984; Acharya et al. 2016). Nonetheless, because observed behavior reflects both preferences and choice constraints, we explore the robustness of our results to their inclusion. As shown in Table A19 (Appendix B.9), estimates for smoking and lending behaviors remain largely unchanged when controlling for income and education. However, for the probability of holding financial assets, the estimated association with the PGI CS disappears once these variables are included. This pattern may indicate that income and education mediate the relationship between genetic endowments and financial behavior. Alternatively, it may reflect collider bias, complicating interpretation. We therefore retain the exclusion of income and education from our main specification.

Selection. Since participation in genotyping within SOEP-G was voluntary, sample selection could introduce bias if individuals who provided saliva samples systematically differ from those who did not. To assess this, we estimate a selection model based on observable demographic characteristics and re-weight our main analyses using inverse probability weights. The resulting estimates, presented in Tables A20 and A21 (Appendix B.10), are consistent with our baseline findings, suggesting that self-selection into genotyping does not materially bias the results.

Additionally, one might hypothesize that individuals consenting to genotyping differ in underlying preferences, such as risk aversion. However, Figure A3 compares the distribution of risk aversion for individuals who provided a saliva sample to those who did not and reveals no systematic differences.

4.3. Nature–Nurture Interactions

This section examines potential gene–environment interactions between genetic endowments for cognitive skills and parental resources. Understanding whether nature and nurture interact is critical for assessing the channels underlying the intergenerational transmission of social and economic inequality. If individuals with favorable genetic endowments disproportionately benefit from enriched environments, such complementarities could amplify initial disparities across generations (*e.g.*, Biroli et al. 2025). Economic theory suggests that both innate endowments and acquired skills can increase the productivity of subsequent investments (Heckman 2006; Cunha and Heckman 2008; Cunha et al. 2010). For instance, parents may adjust their investments based on their children's perceived ability, a behavior that may be more pronounced in higher-resource families. If the relationship between genetic endowments and preferences is stronger in more advantaged environments, this would not only inform the economics of skill formation but also point to pathways through which socio-economic inequality may persist across generations.

To empirically assess these interactions, we estimate models that include interaction terms between the PGI for cognitive skills and an indicator for paternal education, which we use as a proxy for childhood socio-economic environment.[26] Specifically, we distinguish between individuals whose fathers attained at least an intermediate school leaving degree (*Realschule* or higher) and those whose fathers hold only a basic school leaving degree (*Hauptschule*).[27] We focus on paternal education given historically higher male labor force participation, making it a more salient determinant of household resources during childhood. The coefficient on the interaction term captures how the estimated association between PGI CS and preferences varies by paternal education.

Tables 5 (risk preferences and patience) and 6 (social preferences) present the results. For risk aversion, we find evidence of a statistically significant positive interaction: the PGI CS is positively associated with risk aversion only among individuals whose fathers attained a higher educational degree, while the association is statistically insignificant in the lower education group. In contrast, for smoking behavior, which servers as a revealed preference measure of risk tolerance, we find no significant interaction.

For patience, the results suggest no meaningful interaction between PGI CS and paternal education, either for stated patience or for the behavioral measure (holding

[26]Parental education is a strong predictor of socio-economic status in Germany (*e.g.*, Dustmann 2004; Dodin et al. 2024).

[27]In the German system, students were historically tracked into these educational tiers based on academic aptitude assessed in elementary school. See Dustmann (2004) for details.

financial assets). That is, we find no evidence that genetic endowments and parental education are complementary in shaping patience-related outcomes.

TABLE 5. Nature-nurture interactions and risk/time preferences using paternal education

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
PGI CS	-0.020	-0.027**	-0.002	0.036**
	(0.026)	(0.012)	(0.029)	(0.014)
High paternal ed.	-0.005	-0.002	-0.020	0.069**
	(0.058)	(0.026)	(0.063)	(0.030)
High paternal ed. $\times$ PGI CS	0.144***	-0.027	0.017	-0.012
	(0.045)	(0.021)	(0.048)	(0.024)
Observations	2013	1816	1928	1975
R^2	0.146	0.146	0.064	0.102
Test for lin. combination	0.001	0.002	0.705	0.226

Notes: Table 5 displays the association between PGI for cognitive skills and preference measures based on OLS regressions with sample split by paternal education. The variable for interaction is an indicator that equals to 1 if an individual's father had an above-basic education level and zero otherwise. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year. Robust standard errors are included in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Similarly, for most social preference outcomes, including altruism and the frequency of lending belongings, interaction terms are statistically insignificant and small in magnitude. Across all tested domains, the coefficients on the interaction terms are generally indistinguishable from zero at conventional significance levels. Taken together, these findings suggest little evidence of gene–environment complementarities.

However, the pattern points to a potential substitution mechanism: parental resources and genetic endowments may compensate for one another in shaping certain preferences. We observe this substitutability most clearly in preferences related to revealed patience (*i.e.*, financial asset holding), trust, and lending belongings.

TABLE 6. Nature-nurture interactions and social preferences using paternal education

	(1) Trust	(2) Pos. recip.	(3) Neg. recip.	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.072**	-0.040	-0.047	0.025	0.026	0.087***	0.086***
	(0.028)	(0.033)	(0.030)	(0.029)	(0.030)	(0.028)	(0.027)
High paternal ed.	0.165***	-0.041	-0.067	0.061	-0.037	0.137**	0.087
	(0.063)	(0.064)	(0.063)	(0.063)	(0.059)	(0.059)	(0.057)
High paternal ed. $\times$ PGI CS	0.046	0.056	-0.029	0.005	-0.034	-0.017	-0.056
	(0.048)	(0.052)	(0.048)	(0.049)	(0.049)	(0.047)	(0.047)
Observations	1920	1938	1937	1874	1910	1896	1897
R^2	0.115	0.066	0.084	0.117	0.129	0.161	0.222
Test for lin. combination	0.003	0.678	0.044	0.432	0.835	0.067	0.427

Notes: Table 6 displays the association between PGI for cognitive skills and preference outcomes based on OLS regressions with split by paternal education. The variable for interaction is an indicator that equals to 1 if an individual's father had an above-basic education level and zero otherwise. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year. Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

5. Discussion & Conclusion

This study is the first to document a systematic relationship between genetic endowments for cognitive skills and economic preferences. Leveraging a novel representative panel that combines genome-wide data with both stated and revealed preference measures, we find that individuals with stronger genetic predispositions for cognitive skills exhibit distinct patterns in risk, time, and social preferences.

Our contribution is fourfold. First, we offer new insights into how genetic endowments for cognitive skills, measured using a well-validated polygenic index, relate to a broad spectrum of economic preferences. In contrast to earlier studies that rely on contemporaneous cognitive skill measures such as IQ, which reflect both genetic ability and environmental investments, our approach focuses on inherited cognitive endowments. Since investment decisions are themselves likely to correlate with preferences, estimates based on observed cognitive skills may conflate cause and consequence. Leveraging genetic data and comprehensive controls, our estimates more cleanly isolate the association between genetic endowments for cognitive skills and economic preferences.

Second, we show that this genetic variation is systematically related not only to risk, time, and social preferences but also to economically relevant behaviors such as smoking, saving, and prosociality. Third, we test for interactions between genetic endowments and parental resources, shedding light on the role of early-life environments in moderating genetic associations and contributing to the literature on preference formation and the intergenerational transmission of economic inequality. Fourth, we apply Oster (2019)'s sensitivity analysis, which is commonly used in applied microeconomics, to the context of social-science genomics, thereby providing a transparent and systematic robustness assessment in a context where concerns about unobserved confounding are particularly pronounced. We document robust and consistent evidence that genetic endowments for cognitive skills are systematically related to specific economic preferences and behaviors.

A higher PGI for cognitive skills is associated with greater risk aversion, higher levels of trust, and lower negative reciprocity, as well as a lower likelihood of smoking, a higher probability of saving in financial assets, and an increased likelihood of resource sharing. By contrast, we find little evidence of systematic associations with patience, altruism, or helping behavior. These findings suggest that genetic predispositions for cognitive ability have stronger predictive power for some preference domains than others, highlighting the importance of distinguishing between different dimensions of economic preferences.

We subject our estimates to a series of robustness checks. Controlling for detailed family background characteristics and population stratification does not materially affect the results. Sensitivity analyses following Oster (2019) suggest that selection on unobservables would need to be substantially stronger than selection on observables to overturn the estimated associations. These findings support a cautious reduced-form interpretation of the relationship between genetic endowments for cognitive skills and economic preferences.

In addition, tests for gene–environment interactions reveal limited evidence that parental resources systematically amplify or attenuate these associations. On the contrary, we find suggestive evidence of substitution patterns, particularly for revealed patience (*e.g.*, financial asset holding) and for some social preferences. These findings imply that parental investments may, in some domains, offset associations between genetic endowments and economic behavior. This has important implications for policymakers seeking to reduce inequality in economic outcomes rooted in initial endowments.

Our findings differ in part from those of previous studies (*e.g.* Dohmen et al. 2010; Burks et al. 2009; Benjamin et al. 2013), which document lower risk aversion among individuals with higher cognitive test scores. Several factors may account for these differences. First, while prior analyses rely on contemporaneous cognitive test scores that reflect both inherited ability and environmental influences, our approach focuses on genetic endowments fixed at conception. These constitute two different sources of variation: the first encompasses variation due to both genetic endowments *and* environmental factors, such as schooling and extracurricular activities. The latter captures only the variation that can be traced back to genetic endowments. The extent to which the variation due to nurture exerts a causal effect on cognitive skills remains unanswered and is, admittedly, a difficult question.

Second, differences in measurement approaches may also contribute. Most earlier studies[28] rely primarily on incentivized experimental tasks, while our study draws on validated survey-based measures. Taken together, these complementary approaches underscore the importance of both identification strategy and measurement tradition in shaping conclusions about the relationship between cognitive ability and economic preferences (Dohmen et al. 2018).

These findings have broader implications for economics. Preferences are central to

[28]An exception is Dohmen et al. (2010), who also use a stated preference measure to corroborate their experimental findings.

models of savings, labor supply, risk sharing, and cooperation. Our evidence suggests that part of the observed heterogeneity in preferences reflects stable genetic endowments present at conception, specifically endowments related to cognitive skills. This points to a potential channel through which inequality may be transmitted across generations: if both economic preferences and cognitive ability are partly shaped by heritable factors, disparities in economic outcomes may be reinforced over time. At the same time, our null results for several preference domains, along with the limited evidence of gene–environment interactions, suggest that preferences are not rigidly determined by genetic endowments. This leaves scope for environmental influences and targeted interventions to shape economic behavior, even in the presence of underlying biological differences.

More broadly, our findings indicate that heterogeneity in economic preferences is systematic and persistent, and partly associated with genetic variation. Importantly, this does not imply that economic models must become more complex or less tractable. As demonstrated by Koulovatianos et al. (2019), substantial heterogeneity in risk and time preferences can be accommodated within analytically transparent models. In the same spirit, our results help ground such models in empirically observed sources of individual-level variation, clarifying links between micro-level behavior, macroeconomic outcomes, and policy design.

Taken together, our findings suggest that the relationship between genetic endowments and preferences deserves greater attention in economics. Understanding the channels through which preferences are formed and transmitted is essential for analyzing inequality, evaluating the effectiveness of policy interventions, and designing institutions that foster welfare. Future research could build on this work by examining how genetic predispositions interact with policy environments and by exploring whether similar associations hold in other countries or contexts.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen, "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects," *American Political Science Review*, 2016, *110* (3), 512–529.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling," *The Journal of Human Resources*, 2005, *40* (4), 791–821.

Arold, Benjamin W., Paul Hufe, and Marc Stoeckli, "Genetic Endowments, Educational Outcomes and the Mediating Influence of School Quality," *Journal of Political Economy: Microeconomics*, 2023. In press.

Backes-Gellner, Uschi, Holger Herz, Michael Kosfeld, and Yvonne Oswald, "Do Preferences and Biases Predict Life Outcomes? Evidence from Education and Labor Market Entry Decisions," *European Economic Review*, 2021, *134* (103709).

Barban, Nicola, Elisabetta De Cao, and Marco Francesconi, "Gene-Environment Effects on Female Fertility," CEPR Discussion Papers 16603 Oct 2021.

Becker, G. S. and N. Tomes, "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 1986, *4* (3 Pt. 2), 1–47.

Becker, Joel, Casper A. P. Burik, Grant Goldman, Nancy Wang et al., "Resource Profile and User Guide of the Polygenic Index Repository," *Nature Human Behaviour*, 2021, *5* (12), 1744–1758.

Benjamin, Daniel J., David Cesarini, Matthijs J. H. M. van der Loos, Christopher T. Dawes et al., "The Genetic Architecture of Economic and Political Preferences," *Proceedings of the National Academy of Sciences of the United States of America*, May 2012, *109* (21), 8026–8031.

___ , ___ , Patrick Turley, and Alexander Strudwick Young, "Social-Science Genomics: Progress, Challenges, and Future Directions," NBER Working Paper 32404 May 2024.

___ , Sebastian A. Brown, and Jesse M. Shapiro, "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association*, 2013, *11* (6), 1231–1255.

Biroli, Pietro, Titus Galama, Stephanie von Hinke, Hans van Kippersluis et al., "The Economics and Econometrics of Gene–Environment Interplay," *The Review of Economic Studies*, 06 2025, p. rdaf034.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences of the United States of America*, 2009, *106* (19), 7745–7750.

Caliendo, Marco, Frank M Fossen, and Alexander S Kritikos, "Risk Attitudes of Nascent Entrepreneurs-New Evidence from an Experimentally Validated Survey," *Small Business Economics*, 2009, *32* (2), 153–167.

Cesarini, David, Christopher T. Dawes, James H. Fowler, Magnus Johannesson et al., "Heritability of Cooperative Behavior in the Trust Game," *Proceedings of the National Academy of Sciences of the United States of America*, Mar 2008, *105* (10), 3721–3726.

___ , Christopher T Dawes, Magnus Johannesson, Paul Lichtenstein et al., "Genetic Variation in Preferences for Giving and Risk Taking," *The Quarterly Journal of Economics*, 2009, *124* (2), 809–842.

Chen, Chia-Ching, I-Ming Chiu, John Smith, and Tetsuji Yamada, "Too Smart to Be Selfish? Measures of Cognitive Ability, Social Preferences, and Consistency," *Journal of Economic Behavior & Organization*, Jun 2013, *90*, 112–122.

Cobb-Clark, Deborah A., Sarah C. Dahmann, Daniel A. Kamhöfer, and Hannah Schildberg-Hörisch, "Surveillance and Self-Control," *The Economic Journal*, 2023, *134* (660), 1666–1682.

Cobb-Clark, Deborah A., Sarah C. Dahmann, and Nathan Kettlewell, "Depression, Risk Preferences, and Risk-Taking Behavior," *Journal of Human Resources*, 2022, *57* (5), 1566-1604.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, May 2007, *97* (2), 31–47.

___ **and James J Heckman**, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *The Journal of Human Resources*, 2008, *43* (4), 738–782.

___ **, James Heckman, and Susanne Schennach**, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

Davies, Gail, Max Lam, Sarah E. Harris, Joey W. Trampush et al., "Study of 300,486 Individuals Identifies 148 Independent Genetic Loci Influencing General Cognitive Function," *Nature Communications*, 2018, *9* (1), 2098.

Decker, Simon and Hendrik Schmitz, "Health Shocks and Risk Aversion," *Journal of Health Economics*, Dec 2016, *50*, 156–170.

Demange, Perline A, Margherita Malanchini, Travis T Mallard, Pietro Biroli et al., "Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction," *Nature genetics*, 2021, *53* (1), 35–44.

Dodin, Majed, Sebastian Findeisen, Lukas Henkel, Dominik Sachs et al., "Social Mobility in Germany," *Journal of Public Economics*, Apr 2024, *232* (105074), 105074.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Representative Trust and Reciprocity: Prevalence and Determinants," *Economic Inquiry*, 2008, *46* (1), 84–90.

___ **,** ___ **,** ___ **, and** ___ , "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2010, *100* (3), 1238–1260.

___ **,** ___ **,** ___ **, and** ___ , "The Intergenerational Transmission of Risk and Trust Attitudes," *The Review of Economic Studies*, Apr 2012, *79* (2), 645–677.

___ **,** ___ **,** ___ **, and** ___ , "On the Relationship Between Cognitive Ability and Risk Preference," *Journal of Economic Perspectives*, 2018, *32* (2), 115–134.

___ **, David Huffman, Jürgen Schupp, Armin Falk et al.**, "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

Dunn, Olive Jean, "Multiple Comparisons Among Means," *Journal of the American Statistical Association*, 1961, *56* (293), 52–64.

Dustmann, Christian, "Parental Background, Secondary School Track Choice, and Wages," *Oxford Economic Papers*, 2004, *56* (2), 209–230.

Dworkin, Ronald, "What Is Equality? Part 1: Equality of Welfare," *Philosophy & Public Affairs*, 1981, *10* (3), 185–246.

___ , "What Is Equality? Part 2: Equality of Resources," *Philosophy & Public Affairs*, 1981, *10* (4), 283–345.

Elinder, Mikael and Oscar Erixson, "An Inquiry into the Relationship Between Intelligence and Prosocial Behaviour: Evidence from Swedish Population Registers," *Economic Journal*, Nov 2024, pp. 1141–1166.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke et al., "Global Evidence on Economic Preferences," *The Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

—, —, —, David Huffman et al., "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Management Science*, 2022, *69* (4), 1935–1950.

Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia, "Your Place in the World: Relative Income and Global Inequality," *American Economic Journal: Economic Policy*, November 2022, *14* (4), 232–268.

Fehr, Ernst, "On the Economics and Biology of Trust," *Journal of the European Economic Association*, 2009, *7* (2–3), 235–266.

Fey, Gerrit and Renz P Ringsleben, "Aktionärszahlen 2023: Erneut über 12 Millionen. Aktiensparer trotzen der Zinswende," 2024. Published on January 11, 2024. Accessed on September 18, 2025.

Fischbacher, Urs, Levent Neyse, David Richter, and Carsten Schröder, "Adding Household Surveys to the Behavioral Economics Toolbox: Insights from the SOEP Innovation Sample," *Journal of the Economic Science Association*, 2024, *10* (1), 136–151.

Fleurbaey, Marc, *Fairness, Responsibility, and Welfare*, Oxford: Oxford University Press, 2008.

Fraemke, Deniz, Yayouk E. Willems, Aysu Okbay, Ulman Lindenberger et al., "Polygenic Associations With Educational Attainment in East Versus West Germany: Differences Emerge After Reunification," *Psychological Science*, 2025, *36* (7), 559–573.

Fuchs, Victor R., "Time Preference and Health: An Exploratory Study," in "Economic Aspects of Health" NBER Chapters, National Bureau of Economic Research, Inc, January 1982, pp. 93–120.

Fudenberg, Drew and David K. Levine, "A Dual-Self Model of Impulse Control," *American Economic Review*, Dec 2006, *96* (5), 1449–1476.

Genç, Erhan, Caroline Schlüter, Christoph Fraenz, Larissa Arning et al., "Polygenic Scores for Cognitive Abilities and Their Association With Different Aspects of General Intelligence—A Deep Phenotyping Approach," *Molecular Neurobiology*, 2021, *58* (8), 4145–4156.

Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh et al., "The German Socio-Economic Panel (SOEP)," *Journal of Economics and Statistics*, 2019, *239* (2), 345–360.

Heckman, James J., "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, Jun 2006, *312* (5782), 1900–1902.

—, Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

Houmark, Mikkel Aagaard, Victor Ronda, and Michael Rosholm, "The Nurture of Nature and the Nature of Nurture: How Genes and Investments Interact in the Formation of Skills," *American Economic Review*, 2024, *114* (2), 385–425.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *Journal of*

Monetary Economics, 2011, *58* (6-8), 466–477.

Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer, "Worker Beliefs About Outside Options," *The Quarterly Journal of Economics*, 2024, *139* (3), 1505–1556.

Kahneman, Daniel and Shane Frederick, "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in Thomas Gilovich, Dale Griffin, and Daniel Kahneman, eds., *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, 2002, pp. 49–81.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *Journal of Political Economy*, 2018, *126* (3), 937–990.

Koellinger, Philipp D., Aysu Okbay, Hyeokmoon Kweon, Annemarie Schweinert et al., "Cohort Profile: Genetic Data in the German Socio-Economic Panel Innovation Sample (SOEP-G)," *PLOS ONE*, Nov 2023, *18* (11), e0294896.

Koulovatianos, Christos, Carsten Schröder, and Ulrich Schmidt, "Do Demographics Prevent Consumption Aggregates from Reflecting Micro-Level Preferences?," *European Economic Review*, 2019, *111*, 166–190.

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

Lindqvist, Erik and Roine Vestman, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, Jan 2011, *3* (1), 101–128.

List, John A., Azeem M. Shaikh, and Yang Xu, "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, Dec 2019, *22* (4), 773–793.

Lönnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C. Wichardt, "Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison," *Journal of Economic Behavior & Organization*, 2015, *119*, 254–266.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, Feb 2011, *126* (1), 373–416.

Menozzi, P., A. Piazza, and L. Cavalli-Sforza, "Synthetic Maps of Human Gene Frequencies in Europeans: These Maps Indicate That Early Farmers of the Near East Spread to All of Europe in the Neolithic," *Science*, 1978, *201* (4358), 786–792.

Nei, Masatoshi, "Analysis of Gene Diversity in Subdivided Populations," *Proceedings of the National Academy of Sciences*, 1973, *70* (12), 3321–3323.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

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

Rabin, Matthew and Georg Weizsäcker, "Narrow Bracketing and Dominated Choices," *American Economic Review*, 2009, *99* (4), 1508–1543.

Rawls, John, *A Theory of Justice*, Cambridge, MA: Harvard University Press, 1971.

Read, Daniel, George Loewenstein, and Matthew Rabin, "Choice Bracketing," *Journal of Risk and Uncertainty*, 1999, *19* (1–3), 171–197.

Richter, David and Jürgen Schupp, "The SOEP Innovation Sample (SOEP IS)," *Schmollers Jahrbuch*, Sep 2015, *135* (3), 389–399.

Rosenbaum, Paul, "The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Oct 1984, *147* (5), 656–666.

Roth, Christopher and Johannes Wohlfart, "Experienced Inequality and Preferences for Redistribution," *Journal of Public Economics*, Nov 2018, *167*, 251–262.

Serra-Garcia, Marta, "Risk Attitudes and Conflict in the Household," *The Economic Journal*, June 2022, *132* (642), 767–795.

Skoblow, Hanamori F., Christine M. Proulx, Rahmet Akpolat, and Francisco Palermo, "Early-Life Socioeconomic Position and Later-Life Cognitive Functioning: A Meta-Analysis," *Social Science & Medicine*, 2024, *359*, 117267.

Sunde, Uwe, Thomas Dohmen, Benjamin Enke, Armin Falk et al., "Patience and comparative development," *The Review of Economic Studies*, 2022, *89* (5), 2806–2840.

Thaler, Richard H. and H. M. Shefrin, "An Economic Theory of Self-Control," *The Journal of Political Economy*, 1981, *89* (2), 392–406.

Vallentyne, Peter, "Brute Luck, Option Luck, and Equality of Initial Opportunities," *Ethics*, 2002, *112* (3), 529–557.

Vischer, Thomas, Thomas Dohmen, Armin Falk, David Huffman et al., "Validating an Ultra-Short Survey Measure of Patience," *Economics Letters*, 2013, *120* (2), 142–145.

von Hinke, Stephanie, George Davey Smith, Debbie A. Lawlor, Carol Propper et al., "Genetic Markers as Instrumental Variables," *Journal of Health Economics*, Jan 2016, *45*, 131–148.

Wallace, Björn, David Cesarini, Paul Lichtenstein, and Magnus Johannesson, "Heritability of Ultimatum Game Responder Behavior," *Proceedings of the National Academy of Sciences of the United States of America*, October 2007, *104* (40), 15631–15634.

Westfall, P H and S Stanley Young, *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment* Wiley Series in Probability and Statistics, Nashville, TN: John Wiley & Sons, 1992.

Wooldridge, Jeffrey M, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 2010.

Online appendices: Genes, Cognitive Skills, and Preferences

Daniel Graeber, Tanaporn Maneein and Carsten Schroeder

January 17, 2026

Daniel Graeber: SOEP at DIW Berlin, IZA Bonn, and CEPA. **Tanaporn Maneein:** DIW Berlin and University of Potsdam. **Carsten Schröder:** SOEP at DIW Berlin and Freie Universität Berlin. **Corresponding author:** Daniel Graeber, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany. Email: dgraeber[at]diw.de. **Acknowledgments:** We thank Daniel Benjamin, Martin Biewen, Anna Bindler, Marco Caliendo, David Cesarini, Anthony Lepinteur, Katrine V. Løken, Levent Neyse, Thomas Siedler, Uwe Sunde, and participants at the Cluster Seminar on Public Finances and Living Conditions, the European Social Sciences and Genomics Conference 2025, the Annual Conference of the European Society for Population Economics, and the University of Potsdam for helpful comments. **Data availability:** This study uses data from SOEP-IS 2020 (DOI:10.5684/soep.is.2020), which can be accessed through the SOEP Research Data Center. **Ethics:** Data collection for this study received ethical approval from the Vrije Universiteit Amsterdam, School of Business and Economics (Application No. 20181018.1.pkr730), and from the Max Planck Society (Application No. 2019_16). **Declaration of competing interests:** The authors declare no conflict of interest.

Appendix A. Additional data details

Table A1. Factor loadings for trust and reciprocity

Preference measure	Survey question/statement	Factor loading
Trust	When dealing with strangers, it is better to be cautious before trusting them	0.6108
	On the whole, one can trust people	-0.7818
	Nowadays one can't depend on anyone	0.7996
Positive reciprocity	If someone does me a favor, I am prepared to return it	0.7161
	I go out of my way to help somebody who has been kind to me in the past	0.8319
	I am ready to assume personal costs to help somebody who helped me in the past	0.7547
Negative reciprocity	If I suffer a serious wrong, I will take revenge as soon as possible, no matter	0.8720
	If somebody puts me in a difficult position, I will do the same to him/her	0.8827
	If somebody offends me, I will offend him/her back	0.7877

Notes: Responses for trust items range from 1 (Agree completely) to 4 (Disagree completely). Responses for reciprocity items range from 1 (Does not apply to me at all) to 7 (Applies to me perfectly). Higher factor scores indicate higher levels of trust or reciprocity.

Table A2. Definition of variables and response categories

Variable	Definition	Response Categories
Risk aversion	Willingness to take risks (reverse-coded)	0: "Not at all" to 10: "Fully prepared to take risks"
Patience	General level of patience	0: "Very impatient" to 10: "Very patient"
Trust	General trust in others (Factor score based on PCA):	1: "Agree completely" to 4: "Disagree completely"
	– "On the whole, one can trust people."	
	– "Nowadays one can't depend on anyone."	
	– "When dealing with strangers, it is better to be cautious before trusting them."	
Pos. reciprocity	Willingness to return positive actions (Factor score based on PCA):	1: "Does not apply at all" to 7: "Applies completely"
	– "If someone does me a favor, I am prepared to return it."	
	– "I go out of my way to help somebody who has been kind to me in the past."	
	– "I am ready to assume personal costs to help somebody who helped me in the past."	
Neg. reciprocity	Willingness to retaliate against negative actions (Factor score based on PCA):	1: "Does not apply at all" to 7: "Applies completely"

Continued on next page

Table A2 – continued from previous page

Variable	Definition	Response Categories
	– "If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost." – "If somebody puts me in a difficult position, I will do the same to him/her." – "If somebody offends me, I will offend him/her back."	
Altruism	Importance of being there for others (reverse-coded)	1: "Very important" to 4: "Not at all important"

Notes: Table A2 displays outcomes, together with the definition and the response categories.

Table A3. Definition of control variables and response Categories

Variable	Definition	Response Categories
PGI CS	Polygenic index for cognitive skills	Numerical scale, standardized to have mean zero and standard deviation one
Female	Indicator for being female	0: "Male"; 1: "Female"
Year of birth	Year of respondent's birth	Numeric variable
Age	Age, defined as the difference between survey year and year of birth from the year when outcome was observed	Numerical variable
Survey year	Year when interview was conducted	Numerical variable
Principal components of genetic data	First ten principal components of the respondent's genetic data	Numerical scale, standardized to have mean zero and standard deviation one
Paternal education	Father's level of education	1: "Do not know"; 2: "Secondary School"; 3: "Intermediate School"; 4: "Technical School"; 5: "Upper Secondary School"; 6: "Other Degree"; 7: "No School Degree"
Maternal education	Mother's level of education	1: "Do not know"; 2: "Secondary School"; 3: "Intermediate School"; 4: "Technical School"; 5: "Upper Secondary School"; 6: "Other Degree"; 7: "No School Degree"

Continued on next page

Table A3 – continued from previous page

Variable	Definition	Response Categories
Paternal occupational status	Father's occupational status when respondent was 15 years old	1: "Missing"; 2: "Not Employed"; 3: "Self-Employed"; 4: "Blue-Collar"; 5: "White-Collar"; 6: "Civil Service"
Maternal occupational status	Mother's occupational status when respondent was 15 years old	1: "Missing"; 2: "Not Employed"; 3: "Self-Employed"; 4: "Blue-Collar"; 5: "White-Collar"; 6: "Civil Service"
Parents separated	Indicator that is equal to 3 if respondent lived her first 15 years in the same household as biological parents and 2 otherwise	1: "Missing"; 2: "Other types of Household"; 3: "Two-Parents Household"

Notes: Table A3 displays control variables, together with their definitions and response categories.

Table A4. Covariates balance

	PGI CS < Median		PGI CS ≥ Median		
	Mean or N	SD or (%)	Mean or N	SD or (%)	Std. Diff.
Gender					
Male	515	44.8	535	46.6	.0349
Female	634	55.2	614	53.4	
Birth year	1,965	19	1,963	19.3	.12
Father's education					
Upper secondary sch.	127	11.1	208	18.1	.223
Intermediate sch.	186	16.2	167	14.5	
Technical sch.	9	.783	8	.696	
Secondary sch.	636	55.4	617	53.7	
No school degree	26	2.26	29	2.52	
Other degree	26	2.26	20	1.74	
Missing	139	12.1	100	8.7	
Mother's education					
Upper secondary sch.	71	6.18	124	10.8	.198
Intermediate sch.	232	20.2	236	20.5	
Technical sch.	8	.696	14	1.22	
Secondary sch.	680	59.2	645	56.1	
No school degree	41	3.57	34	2.96	
Other degree	19	1.65	24	2.09	
Missing	98	8.53	72	6.27	
Father's occupation					
Self-employed	114	9.92	155	13.5	.229
White-collar	333	29	344	29.9	
Civil service	66	5.74	112	9.75	
Blue-collar	434	37.8	365	31.8	
Not employed	103	8.96	104	9.05	
Missing	99	8.62	69	6.01	
Mother's occupation					
Self-employed	61	5.31	74	6.44	.142
White-collar	354	30.8	400	34.8	
Civil service	16	1.39	25	2.18	
Blue-collar	235	20.5	190	16.5	
Not employed	372	32.4	358	31.2	
Missing	111	9.66	102	8.88	
HH. env. growing up					
Two-parents HH.	836	72.8	838	72.9	.0668
Other types of HH.	264	23	247	21.5	
Missing	49	4.26	64	5.57	

Notes: Table A4 displays the standardized differences in covariates between individuals with polygenic index for cognitive skills above and below the median. A standardized difference greater than 0.1 in absolute value could indicate covariate imbalance between the two groups, as proposed by Austin (2001).
Source: Authors' calculations based on SOEP-IS 2020.

Appendix B. Details on the sensitivity analyses

B.1. Replication of previous studies

Our baseline finding on risk aversion diverges markedly from previous results in the literature (Dohmen et al. 2010; Benjamin, Brown, and Shapiro 2013; Burks et al. 2009). We interpret this divergence as likely driven by differences in how cognitive skills are measured. In particular, we argue that our proposed measure captures a distinct dimension of systematic variation that may not be fully accounted for by commonly used proxies such as test scores, in which endowment-related variation may be confounded with postnatal investment. To support this interpretation, we conduct a supplementary analysis[1] that follows previous studies in using standardized test scores as a proxy for cognitive skills.

Table A5 reports the results. When test scores are used, we are able to replicate the negative relationship between cognitive skills and risk aversion documented in previous studies, supporting our interpretation that the opposing direction of the baseline result stems from differences in measurement approaches. We additionally find that higher cognitive skills are associated with greater prosociality, in line with existing evidence (see, e.g., Elinder and Erixson 2024). We find no statistically significant relationship between cognitive skills and stated patience.

B.2. Relative importance of selection on unobservables consistent with a null effect

The key insight of Oster's delta method is that the degree of selection on observables can inform the degree of selection on unobservables. The proportionality factor δ is defined as:

$$\delta = \frac{\sigma_{PGI\ CS,U}}{\sigma_{PGI\ CS,O}}, \tag{A1}$$

where $\sigma_{PGI\ CS,U}$ represents the covariance between the treatment and unobserved confounders, U, and $\sigma_{PGI\ CS,O}$ represents the covariance between the treatment and observed controls, O, which in our case contains our controls for age, year of birth, parental controls and genetic PCs.

[1] This analysis uses the SOEP-Core sample, from which the SOEP-IS and SOEP-G subsamples are drawn and contains standardized test scores. Following Dohmen et al. (2010), we focus on data around 2006.

TABLE A5. Association between cognitive skills and stated preferences

	(1) Trust	(2) Pos. Reciprocity	(3) Neg. Reciprocity	(4) Altruism	(5) Patience	(6) Risk Aversion
Cognitive skills	0.106***	0.025	-0.074***	0.043***	0.013	-0.074***
	(0.017)	(0.018)	(0.018)	(0.016)	(0.017)	(0.015)
Observations	4507	3631	3619	4513	4513	5541
R^2	0.025	0.021	0.065	0.048	0.022	0.101

Notes: Table A5 displays the partial correlations between economic preferences and cognitive skills controlling for year of birth and sex. Economic preferences include self-reported levels of trust, altruism, positive reciprocity, negative reciprocity, patience, and risk aversion. Cognitive test scores correspond to the standardized average of standardized scores in the animal naming and numerical task in 2006. Standardized risk aversion was measured in 2006. Positive and negative reciprocity have been measured in 2005. Social trust, altruism, and patience have been measured in 2008. Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01.

To assess robustness, Oster (2019) derives the bias-adjusted treatment effect:

$$\text{(A2)}\quad \beta_1^* = \beta_{1,\text{controlled}} - \delta\left(\beta_{1,\text{uncontrolled}} - \beta_{1,\text{controlled}}\right) \times \frac{R_{\text{max}} - R_{\text{controlled}}}{R_{\text{controlled}} - R_{\text{uncontrolled}}}.$$

Here, R_{max} represents the hypothetical R^2 if all confounders were observed, $R_{\text{controlled}}$ is the R^2 with observed controls, and $R_{\text{uncontrolled}}$ is the R^2 without controls. Typically, R_{max} is set to $1.3 \cdot R_{controlled}$ (Oster 2019). The intuition for this recommendation is that, data in social science is gathered to explain the phenomena we are interested in. As a result, we would not expect that many controls are missing. Further, most social science data comprise measurement errors (Oster 2019). In order to understand the selection based on unobservables necessary to be consistent with the absence of no effect, *i.e.*, $\beta_1^* = 0$, we set Equation A2 equal to zero and solve for δ. The results are displayed in Table A6 and A7.

Table A6. Relative importance of selection on unobservables - risk and time preferences

	Risk aversion		Smoking		Patience		Fin. assets	
	(1) Base	(2) Full	(3) Base	(4) Full	(5) Base	(6) Full	(7) Base	(8) Full
PGI CS	0.058***	0.050**	-0.048***	-0.035***	0.008	0.003	0.049***	0.034***
	(0.020)	(0.021)	(0.009)	(0.009)	(0.021)	(0.023)	(0.010)	(0.011)
Delta		10.662		5.399		1.987		4.615
Observations	2298	2298	2057	2057	2200	2200	2256	2256
R^2	0.003	0.138	0.013	0.143	0.000	0.061	0.010	0.113

Notes: Table A6 displays the relative importance of selection on unobservables, relative to observables, needed to fully explain the estimates in our main specification, detailed in Equation 4. The outcomes correspond to our set of risk and time preference measures. The odd columns display the effect estimate of a regression of the outcome on the PGI CS. The even columns display the results of our main specification. Delta describes the relative importance of selection on unobservables, relative to selection on observables, to fully explain our effect. Consistent with the recommendation by Oster (2019), we set the R^2 associated with a model, that includes all unobservables, equal to 1.3 times the R^2 in the corresponding even column. Oster (2019) suggests that deltas larger than one can be considered robust. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A7. Relative importance of selection on unobservables - social preferences

	Trust		Pos. reciprocity		Neg. reciprocity		Altruism		Lending help		Lending blongings		Lending money	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Base	Full	Base	Full	Base	Full	Base	Full	Base	Full	Base	Full	Base	Full
PGI CS	0.131***	0.092***	-0.030	-0.025	-0.066***	-0.058***	0.011	0.016	-0.022	-0.009	0.049**	0.053**	0.037*	0.055***
	(0.021)	(0.022)	(0.023)	(0.024)	(0.021)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)
Delta		4.432		10.330		10.823		-13.770		1.889		152.571		-12.498
Observations	2189	2189	2197	2197	2197	2197	2130	2130	2173	2173	2166	2166	2167	2167
R^2	0.017	0.118	0.001	0.053	0.004	0.079	0.000	0.104	0.000	0.130	0.002	0.140	0.001	0.211

Notes: Table A7 displays the relative importance of selection on unobservables, relative to observables, needed to fully explain the estimates in our main specification, detailed in Equation 4. The outcomes correspond to our set of social preference measures. The odd columns display the effect estimate of a regression of the outcome on the PGI CS. The even columns display the results of our main specification. Delta describes the relative importance of selection on unobservables, relative to selection on observables, to fully explain our effect. Consistent with the recommendation by Oster (2019), we set the R^2 associated with a model, that includes all unobservables, equal to 1.3 times the R^2 in the corresponding even column. Oster (2019) suggests that deltas larger than one can be considered robust. Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

B.3. Accounting for individual level heterogeneity

In our baseline specification, we estimate the model using ordinary least squares (OLS) as follows:

$$y_{it} = \beta_0 + \beta_1 PGI\ CS_{it} + \beta_2 X_{it} + \beta_3 Z_{it} + \beta_4 PC_{it} + \epsilon_{it}. \tag{A3}$$

This model assumes that the error term ϵ_{it} captures all unobserved factors affecting the outcome y_{it}.

To address potential unobserved, time-invariant heterogeneity that might influence the outcome but is not captured by our observed covariates, we augment our model by introducing a random effect. Specifically, we allow for an individual-specific component, u_i, which captures persistent characteristics unique to each individual. The augmented model can be written as:

$$y_{it} = \alpha_0 + \alpha_1 PGI\ CS_{it} + \alpha_2 X_{it} + \alpha_3 Z_{it} + \alpha_4 PC_{it} + u_i + \mu_{it}, \tag{A4}$$

where:

- u_i is the individual-specific random effect, which we assume to be a random draw from a normal distribution, that is,

  $$u_i \sim \mathcal{N}(0, \sigma_u^2),$$

  capturing all time-invariant unobserved heterogeneity;
- ϵ_{it} is the idiosyncratic error term.

By specifying the model in this way, we allow the intercept to vary across individuals. The individual-specific effect u_i captures unobserved factors such as family background, ancestry effects or even genetic endowments of the parents, among others. This random effects model thus helps to mitigate bias due to omitted time-invariant variables.

The efficiency of the random effects estimator relies on the assumption that the individual-specific effects u_i are uncorrelated with the explanatory variables. Our sensitivity analysis involves estimating the random effects model in Equation (A4) and comparing the resulting coefficient estimates, particularly for β_1, the coefficient on the PGI for cognitive skills, with those from the OLS specification in Equation (A3). If the estimates remain stable across these specifications, it provides evidence that unobserved heterogeneity does not bias our empirical findings.

Table A8. Random effects model for risk/time preferences

	(1) Risk aversion	(2) Smoking	(3) Patience
PGI CS	0.051***	-0.034***	0.000
	(0.015)	(0.009)	(0.020)
Observations	17319	2904	3665
Clusters	2298	2057	2200

Notes: Table A17 displays the association between PGI score for cognitive skills and risk/time preference measures based on random effects model. The results for the probability of holding financial assets are omitted due to the lack of panel data for the outcome. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year. Cluster robust standard errors are included in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A9. Random effects model for social preferences

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.098***	-0.017	-0.065***	0.007	-0.028	0.052**	0.056***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.017)	(0.021)	(0.020)
Observations	3768	4295	4297	3275	6653	2423	2421
Clusters	2194	2200	2200	2130	2173	2166	2167

Notes: Table A9 displays the association between PGI score for cognitive skills and social preference outcomes based on random effects model. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year . Cluster robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

B.4. Addressing pleiotropy through sensitivity analysis with PGIs for non-cognitive skills

Genetic variants may influence multiple phenotypes, a phenomenon known as *pleiotropy* (Benjamin et al. 2024). In our context, pleiotropy implies that some SNPs contributing to the PGI CS may also affect non-cognitive traits. This overlapping genetic influence can potentially confound the estimated association between the PGI CS and the outcome of interest if part of the observed effect is actually driven by genetic influences on other correlated traits.

Furthermore, to the extent that the SNPs informing the PGIs for non-cognitive skills are in linkage disequilibrium with those contributing to the PGI CS, the inclusion of these additional PGIs also helps to account for any correlation arising from shared genetic architecture. In essence, our sensitivity analysis not only adjusts for potential pleiotropic influences but also mitigates any bias introduced by linkage disequilibrium between the SNPs underlying different PGIs.

To address these concerns, we implement a sensitivity analysis by including PGIs for non-cognitive traits, specifically, openness to experience, extraversion, and neuroticism, in our main empirical model. These traits are selected because they capture key dimensions of personality that are commonly associated with both cognitive skills and social outcomes and are included in the SOEP-G (Koellinger et al. 2023). By controlling for these additional PGIs, we aim to isolate the unique effect of the PGI CS, ensuring that the observed association is not spuriously driven by pleiotropic effects.

Tables A10 and A11 display the results from models that incorporate these non-cognitive PGIs.

Table A10. Addressing pleiotropy in the association between PGI CS and risk/time preferences

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
PGI CS	0.053***	-0.032***	-0.006	0.034***
	(0.021)	(0.010)	(0.023)	(0.011)
Observations	2298	2057	2200	2256
R^2	0.143	0.146	0.064	0.114

Notes: Table A10 displays the association between PGI score for cognitive skills and risk/time preference outcomes after controlling for PGIs related to non-cognitive traits to address pleiotropy concerns. In addition to the full specification, single-trait PGIs for extraversion, neuroticism, and openness, dimensions of the Big Five personality traits for which genetic data is available, are included. The full specification controls include age, age squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender, and birth year. Robust standard errors are provided in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A11. Addressing pleiotropy in the association between PGI CS and social preferences

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.089***	-0.025	-0.051**	0.018	-0.007	0.048**	0.051**
	(0.022)	(0.025)	(0.022)	(0.023)	(0.022)	(0.022)	(0.020)
Observations	2189	2197	2197	2130	2173	2166	2167
R^2	0.120	0.053	0.082	0.106	0.131	0.144	0.212

Notes: Table A11 displays the association between PGI score for cognitive skills and social preference outcomes after controlling for PGIs related to non-cognitive traits to address pleiotropy concerns. In addition to the full specification, single-trait PGIs for extraversion, neuroticism, and openness, dimensions of the Big Five personality traits for which genetic data is available, are included. The full specification controls include age, age squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender, and birth year. Robust standard errors are provided in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

B.5. Multiple hypothesis testing adjustments

In our empirical application, multiple hypothesis testing presents a challenge in maintaining statistical validity. When testing multiple hypotheses, the probability of false positives increases, necessitating adjustments to p-values. In this section we apply three widely used procedures: the Bonferroni adjustment (Dunn 1961), Westfall and Young (1992)'s resampling-based approach, and the adjustment proposed by List, Shaikh, and Xu (2019).

Bonferroni adjustment. The Bonferroni adjustment is one of the most widely used methods to control the family-wise error rate (FWER). It achieves this by setting a more stringent significance threshold:

$$p^{\text{adjusted}} = \min(1, m \cdot p) \tag{A5}$$

where m is the number of hypotheses tested, and p is the unadjusted p-value. While conservative, this method ensures strong control of Type I error at the cost of increased Type II errors.

Westfall and Young (1992)'s resampling-based approach. Westfall and Young (1992) introduce a resampling-based method that accounts for the dependency structure among tests. Instead of applying a fixed correction factor, they use permutation techniques to estimate the joint distribution of p-values, allowing for a less conservative adjustment while still controlling the FWER. This method is particularly useful when test statistics are correlated, reducing the loss of statistical power compared to the Bonferroni adjustment.

List, Shaikh, and Xu (2019)'s adjustment for experimental economics. List, Shaikh, and Xu (2019) propose an approach, mainly tailored to experimental economics, that balances statistical power and error control. Their method builds upon family-wise and false discovery rate (FDR) adjustments, recognizing that in experimental economics, researchers often face a large number of simultaneous tests. Their approach suggests a combination of stepwise adjustments and economic theory-based pre-registration to mitigate concerns about data mining.

Comparison of methods. These three methods offer different trade-offs:

- **The Bonferroni adjusment** is simple and ensures strong control over false positives but is overly conservative when hypotheses are correlated.
- **Westfall and Young (1992)** provide an improved alternative that maintains control over false positives while increasing statistical power through resampling techniques.
- **List, Shaikh, and Xu (2019)** offer a more flexible approach for experimental economics, allowing for more nuanced adjustments based on empirical settings and research design.

We apply the these procedures to our hypotheses testing for different groups of hypotheses. The results are displayed in Table A12 and A13.

Table A12. Multiple hypothesis testing

Hyp. num.	Dep. var.	Coeff.	P-value	List et al.	Bonferroni	Westfall-Young	Randomi-zation
1	Risk aversion	0.050	0.016	0.070	0.163	0.086	0.014
2	Smoking	-0.035	0.000	0.000	0.002	0.003	0.000
3	Patience	0.003	0.880	0.883	1.000	0.899	0.879
4	Fin. assets	0.034	0.002	0.034	0.044	0.016	0.003
5	Trust	0.092	0.000	0.000	0.002	0.000	0.000
6	Pos. reciprocity	-0.025	0.298	0.759	1.000	0.759	0.298
7	Neg. reciprocity	-0.058	0.008	0.061	0.101	0.058	0.010
8	Altruism	0.016	0.468	0.859	1.000	0.838	0.465
9	Lending help	-0.009	0.693	0.898	1.000	0.899	0.703
10	Lending belongings	0.053	0.017	0.082	0.158	0.086	0.016
11	Lending Money	0.055	0.008	0.057	0.084	0.058	0.006
All				0.002			

Notes: Table A12 displays the original coeffients and p-values of PGI CS and the adjusted p-values from different methods under different outcomes. The adjustment procedure by List, Shaikh, and Xu (2019) and Westfall and Young (1992) allows for dependencies across outcomes. Bonferroni method restricts an upper bound to the adjusted p-values at 1 (Dunn 1961). Randomization is based on randomised inference as in Young (2018). Statistical inference is based on randomised inference as in Young (2018), in which a multiple hypothesis testing across all equations is performed. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A13. Multiple hypothesis testing by preference domains

Hyp. num.	Dep. var.	Coeff.	P-value	List et al.	Bonferroni	Westfall-Young	Randomi-zation
Family 1: Risk preferences							
1	Risk aversion	0.050	0.016	0.015	0.030	0.016	0.014
2	Smoking	-0.035	0.000	0.000	0.000	0.001	0.000
All			0.000				
Family 2: Time preferences							
1	Patience	0.003	0.880	0.882	1.000	0.881	0.879
2	Fin. assets	0.034	0.002	0.003	0.003	0.003	0.002
All			0.005				
Family 3: Social preferences							
1	Trust	0.092	0.000	0.000	0.001	0.000	0.000
2	Pos. reciprocity	-0.025	0.298	0.644	1.000	0.661	0.298
3	Neg. reciprocity	-0.058	0.008	0.043	0.063	0.043	0.010
4	Altruism	0.016	0.468	0.707	1.000	0.706	0.465
5	Lending help	-0.009	0.693	0.696	1.000	0.706	0.703
6	Lending belongings	0.053	0.017	0.066	0.119	0.065	0.016
7	Lending money	0.055	0.008	0.048	0.059	0.043	0.006
All			0.000				

Notes: Table A13 displays the original coeffients and p-values of PGI CS and the adjusted p-values from different methods under different outcomes according to preference domains. The adjustment procedure by List, Shaikh, and Xu (2019) and Westfall and Young (1992) allows for dependencies across outcomes. Bonferroni method restricts an upper bound to the adjusted p-values at 1 (Dunn 1961). Statistical inference is based on randomized inference as in Young (2018), in which a multiple hypothesis testing across all equations is performed. *Source:* Authors' calculations based on SOEP-IS 2020.

B.6. Quality control procedures: mild-QC vs. strict-QC

In the SOEP-G, two distinct quality control (QC) pipelines were implemented for processing the genetic data:[2]

- **Mild-QC:** This pipeline applies less stringent criteria. It excludes individuals with extreme values, such as those with per-chromosome missingness greater than 50% and marked heterozygosity/homozygosity outliers, while retaining a larger sample size.
- **Strict-QC:** In contrast, the strict-QC pipeline enforces rigorous thresholds by excluding individuals with per-chromosome missingness over 20% and those with excessive heterozygosity/homozygosity. This results in a smaller, but arguably higher-quality dataset.

Despite the stricter criteria, both pipelines yield PGIs with nearly equivalent predictive accuracy (Koellinger et al. 2023). The usage of both quality standards comes with advantages and disadvantages:

- Using the **mild-QC** data when maximizing sample size (and thus statistical power) is essential, especially for analyses that rely on aggregated PGIs.
- Opting for the **strict-QC** data for studies that require very high genotype fidelity, such as analyses of single genetic variants or detailed genotype-phenotype associations.

In our main analyses, we relied on the sample that results from the mild-QC. In Table A14 and A15, we present results based on the sample that results from the strict-QC.

[2]See (Koellinger et al. 2023) for details.

Table A14. Association between PGI CS and risk/time preferences under strict quality control

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
PGI CS	0.045**	-0.027***	-0.001	0.029***
	(0.021)	(0.010)	(0.024)	(0.011)
Observations	2092	1867	2001	2054
R^2	0.136	0.150	0.065	0.112

Notes: Table A14 displays the results for risk/time preferences upon imposing strict quality control on the PGI score for cognitive skills. The associations shown are obtained from regressions with full specification, controlling for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year .Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A15. Association between PGI CS and social preferences under strict quality control

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.091***	-0.029	-0.055**	0.015	-0.012	0.039*	0.051**
	(0.023)	(0.025)	(0.023)	(0.024)	(0.023)	(0.023)	(0.021)
Observations	1991	1998	1998	1937	1975	1972	1973
R^2	0.117	0.064	0.077	0.106	0.132	0.145	0.222

Notes: Table A15 displays the results for social preferences upon imposing strict quality control on the PGI score for cognitive skills. The associations shown are obtained from regressions with full specification, controlling for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year .Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

B.7. Motivation and procedure for controlling measurement error in PGI regressions

PGIs are constructed as weighted sums of allele counts and serve as predictors for various phenotypes. However, because the weights are estimated from finite-sample GWAS summary statistics, a PGI is an *unbiased but noisy* measure of the true *additive SNP factor*, the optimal linear predictor based on all measured SNPs. This noise induces *errors-in-variables bias* in regression analyses that use the PGI as an explanatory variable, typically attenuating estimated coefficients toward zero.

To address this issue, the authors derive a measurement-error correction (also known as regression disattenuation) estimator. The key steps are:

(a) **Characterization of measurement Error:** The PGI is shown to be equivalent to the true additive SNP factor plus classical measurement error. In particular, if we denote the true factor as g_i and the PGI as $\hat{g}_i$, then

$$\hat{g}_i = g_i + e_i,$$

where e_i is mean-zero error that is uncorrelated with g_i. The magnitude of the measurement error is captured by a factor

$$\rho = \sqrt{\frac{h^2_{\text{SNP}}}{R^2}},$$

with h^2_{SNP} representing the SNP heritability (*i.e.*, the maximum predictive power of g_i) and R^2 being the actual variance explained by the PGI.

(b) **Bias in standard regressions:** When the PGI is used as a proxy for the unobserved g_i in a regression model, the estimated coefficients are biased due to measurement error. In the simplest univariate case, the estimated slope is attenuated by a factor of $1/\rho$.

(c) **Corrected estimator:** Becker et al. (2021) extend the standard errors-in-variables framework to the multivariate context. They derive an estimator that "inverts" the attenuation bias by adjusting the regression coefficients using estimates of the variance-covariance structure of the variables and the measurement error parameter ρ. The essential idea is to replace the naive OLS estimates with corrected coefficients that reflect the true effect of the standardized additive SNP factor.

(d) **Practical implementation:** In practice, the procedure involves:

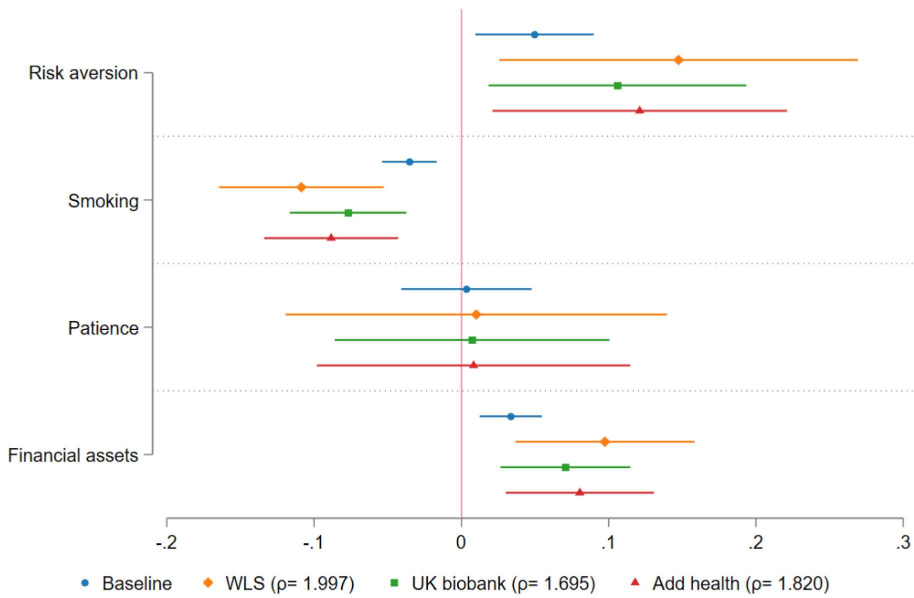
- Estimating h^2_{SNP} (using methods such as Genome-based restricted maximum likelihood (GREML) or BOLT-REML) and R^2 from regressions of the phenotype on the PGI.
- Computing ρ, and then using it to adjust the estimated regression coefficients and standard errors.
- In the univariate case, this amounts to multiplying the OLS coefficient and its standard error by ρ, while in the multivariate case, the adjustment is applied via a matrix correction.

The corrected coefficients are reported in units of the true additive SNP factor, providing more interpretable and comparable effect sizes across studies and different PGIs. This procedure also enhances the validity of empirical inferences by mitigating the downward bias caused by measurement error in the PGI.

In order to understand the extent of attenuation bias in our main results, we apply the measurement error correction to our main estimates. Ideally, we would like to have access to h^2_{SNP} and R^2 from a GWAS in SOEP-G. In absence of these output, we rely on available inputs from the Wisconsin Longitudinal Study (WLS), UK Biobank, and Add Health. Input from the WLS and UK Biobank comes from Becker et al. (2021). The inputs from Add Health are based on Sanz-de Galdeano and Terskaya (2023).[3] The results are displayed in Figure A1 and A2 as well as Table A16. In each of these, we show the estimate based on our main analysis together with the ones which are based on the measurement error corrected versions.

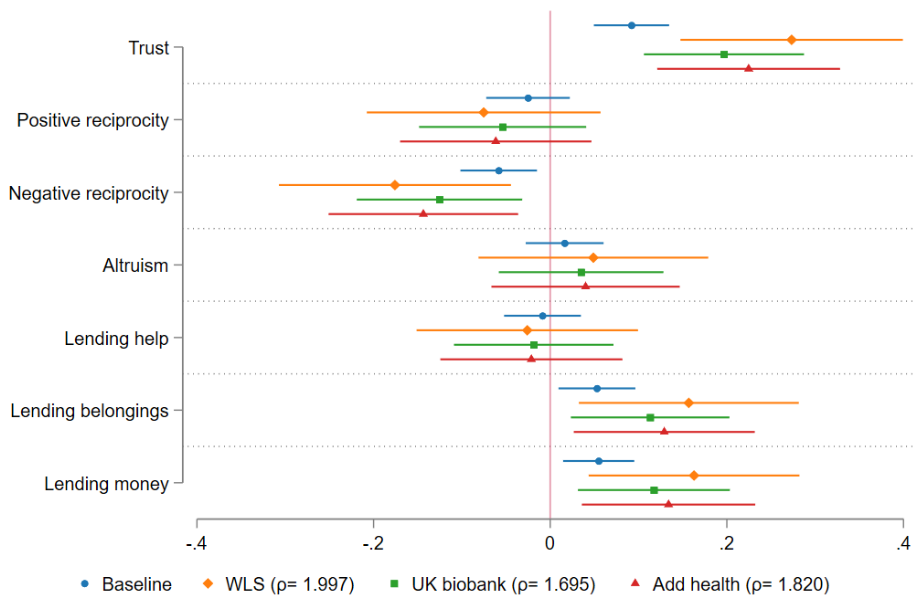
[3]The Python code for the measurement error correction is provided by Becker et al. (2021) via https://github.com/JonJala/pgi_correct. This version is associated with the commit hash `12d46d7`.

FIGURE A1. The effect of the PGI CS on risk aversion and patience - correcting for measurement error) (Becker et al. 2021)



Notes: Figure A1 illustrates the association between PGI CS and risk aversion as well as patience based on our main specification, alongside the corresponding associations using the measurement error correction. For each outcome, results are shown for h^2_{SNP} and R^2 from the WLS, UK Biobank, and Add Health (Becker et al. 2021; Sanz-de Galdeano and Terskaya 2023). In all specifications, controls include gender, age, year of birth, parental school leaving degrees, parental occupational status, an indicator for parental separation, and the first ten genetic principal components. The measurement-error-corrected analyses employ the method of Becker et al. (2021). Throughout, 95% confidence intervals are reported based on analytic, robust standard errors.

Figure A2. The effect of the PGI CS on social preferences - correcting for measurement error(Becker et al. 2021)



Notes: Figure A1 illustrates the association between PGI CS and social preferences based on our main specification, alongside the corresponding associations using the measurement error correction. For each outcome, results are shown for h^2_{SNP} and R^2 from the WLS, UK Biobank, and Add Health (Becker et al. 2021; Sanz-de Galdeano and Terskaya 2023). In all specifications, controls include gender, age, year of birth, parental school leaving degrees, parental occupational status, an indicator for parental separation, and the first ten genetic principal components. The measurement-error-corrected analyses employ the method of Becker et al. (2021). Throughout, 95% confidence intervals are reported based on analytic, robust standard errors.

Table A16. Measurement error correction

	Main results ρ = 0	WLS ρ = 1.997	Add health ρ = 1.820	UK biobank ρ = 1.695
Risk aversion	0.050**	0.147**	0.121**	0.106**
	(0.021)	(0.062)	(0.051)	(0.045)
Smoking	-0.035***	-0.109***	-0.088***	-0.077***
	(0.009)	(0.029)	(0.023)	(0.020)
Patience	0.003	0.010	0.008	0.007
	(0.023)	(0.066)	(0.054)	(0.048)
Fin. assets	0.034***	0.097***	0.080***	0.071***
	(0.011)	(0.031)	(0.026)	(0.023)
Trust	0.092***	0.273***	0.225***	0.197***
	(0.022)	(0.064)	(0.053)	(0.046)
Pos. reciprocity	-0.025	-0.075	-0.062	-0.054
	(0.024)	(0.068)	(0.055)	(0.048)
Neg. reciprocity	-0.058***	-0.176***	-0.144***	-0.125***
	(0.022)	(0.067)	(0.055)	(0.048)
Altruism	0.016	0.049	0.040	0.035
	(0.022)	(0.066)	(0.054)	(0.048)
Lending help	-0.009	-0.026	-0.021	-0.019
	(0.022)	(0.064)	(0.053)	(0.046)
Lending belongings	0.053**	0.157**	0.129**	0.113**
	(0.022)	(0.064)	(0.052)	(0.046)
Lending money	0.055***	0.163***	0.134***	0.117***
	(0.021)	(0.061)	(0.050)	(0.044)

Notes: Table A16 displays the original coefficients PGI CS coefficients preference outcomes and the corresponding measurement-error adjusted statistics computed using the method by Becker et al. (2021). Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

B.8. Checking for non-linearities

Table A17. Non-Linearity between PGI CS and risk/time preferences

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
PGI CS	0.051**	-0.036***	0.002	0.033***
	(0.021)	(0.010)	(0.023)	(0.011)
PGI CS squared	-0.010	0.008	0.014	0.001
	(0.014)	(0.006)	(0.015)	(0.007)
Observations	2298	2057	2200	2256
R^2	0.139	0.144	0.061	0.113

Notes: Table A17 displays the results from examining linearity in the association between PGI score for cognitive skills and risk/time preference measures based on OLS regressions. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender, birth year, Robust standard errors are included in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A18. Non-Linearity between PGI CS and social preferences

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.089***	-0.024	-0.060***	0.018	-0.010	0.050**	0.054***
	(0.022)	(0.025)	(0.023)	(0.023)	(0.022)	(0.022)	(0.020)
PGI CS squared	0.029**	-0.011	0.012	-0.011	0.013	0.026*	0.005
	(0.014)	(0.018)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)
Observations	2189	2197	2197	2130	2173	2166	2167
R^2	0.120	0.053	0.079	0.104	0.130	0.142	0.211

Notes: Table A18 displays the results from examining linearity in the association between PGI score for cognitive skills and social preference outcomes based on OLS regressions. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender, birth year .Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

B.9. Additional controls for revealed preferences outcomes

Table A19. Association between PGI CS and revealed preferences controlling for education and income

	(1) Smoking	(2) Fin. assets	(3) Lending help	(4) Lending belongings	(5) Lending money
Panel A: Main specification controls					
PGI CS	-0.035***	0.034***	-0.009	0.053**	0.055***
	(0.009)	(0.011)	(0.022)	(0.022)	(0.021)
R^2	0.143	0.113	0.130	0.140	0.211
Adjusted R^2	0.092	0.067	0.082	0.093	0.168
Panel B: Main specification and education controls					
PGI CS	-0.023**	0.022*	-0.004	0.038*	0.046**
	(0.010)	(0.011)	(0.022)	(0.022)	(0.021)
R^2	0.154	0.125	0.128	0.150	0.208
Adjusted R^2	0.101	0.077	0.079	0.101	0.163
Panel C: Main specification and income controls					
PGI CS	-0.031***	0.023**	-0.012	0.057**	0.060***
	(0.010)	(0.011)	(0.023)	(0.022)	(0.021)
R^2	0.169	0.154	0.135	0.147	0.207
Adjusted R^2	0.118	0.108	0.086	0.098	0.161
Panel D: Main specification and both additional controls					
PGI CS	-0.022**	0.015	-0.007	0.045**	0.053**
	(0.010)	(0.011)	(0.023)	(0.023)	(0.021)
Observations	1938	2144	2079	2050	2050
R^2	0.170	0.158	0.134	0.154	0.203
Adjusted R^2	0.116	0.110	0.083	0.104	0.156

Notes: Table A19 presents the association of the PGI score for cognitive skills on revealed preferences. The main specification controls in Panel A include gender, year of birth, age, age squared, the first ten principal components of the genetic data, parental educational attainment, occupational status, and an indicator for parental divorce. Panel B adds own education to the control set. Panel C adds log household net income to the main specification. Panel D combines all these controls. Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Source:* Authors' calculations based on SOEP-IS 2020.

B.10. Weighting for selective consent and adjusting for sample selectivity

In many empirical studies, especially those involving sensitive information such as genetic data, only a subset of individuals consent to participate. If the decision to consent is correlated with characteristics related to the outcome, the analysis sample may not be representative of the target population. This *selective consent* can introduce bias into the estimates if not properly addressed.

Estimation of adjustment weights. To correct for selective consent, we estimate adjustment weights that account for differences between those who consent and those who do not. The procedure is as follows:

(a) **Selection model:** We estimate a selection model (using a logistic regression) where the dependent variable is an indicator S_i that equals 1 if an individual is included in the analysis sample and 0 otherwise.

(b) **Predictors:** The model includes a broad set of covariates, W_i, that are believed to influence the likelihood of consent. These covariates may include demographic characteristics, socioeconomic indicators, and other relevant attributes.

(c) **Adjustment factor:** For each observation, we predict the probability $\hat{p}_i = P(S_i = 1 \mid W_i)$ of being included in the analysis sample. The adjustment factor is then given by the inverse of this probability, $\frac{1}{\hat{p}_i}$.

    This procedure extrapolates the analysis sample to represent the population of interest more accurately.

Caveats on covariate selection. A key consideration in this procedure is the choice of covariates W_i used in the selection model. In particular, one must be cautious about including covariates that are themselves outcomes (or post-treatment variables) of the explanatory variable in the primary analysis. If a covariate is affected by the treatment or explanatory variable, controlling for it in the selection model can lead to *post-treatment bias*. Such bias occurs because these variables lie on the causal pathway between the treatment and the outcome, and adjusting for them may remove part of the true effect or even induce spurious associations (Imbens and Rubin 2015; Angrist and Pischke 2008; Morgan and Winship 2014).

Thus, it is crucial that the covariates included in the selection model satisfy the following:

- They should predict participation (or consent) independently of the causal pathway between the primary explanatory variable and the outcome.
- They should be exogenous to the treatment effect – *i.e.*, not be influenced by the explanatory variable under study.

Failure to adhere to these principles may result in adjustment factors that inadvertently control for part of the effect one aims to measure, thereby distorting the estimated impact.

Table A20. Weighted association between PGI CS and risk/time preferences

	(1) Risk aversion	(2) Smoking	(3) Patience	(4) Fin. assets
PGI CS	0.048**	-0.035***	0.004	0.032***
	(0.021)	(0.010)	(0.023)	(0.011)
Observations	2298	2057	2200	2256
R^2	0.139	0.142	0.063	0.115

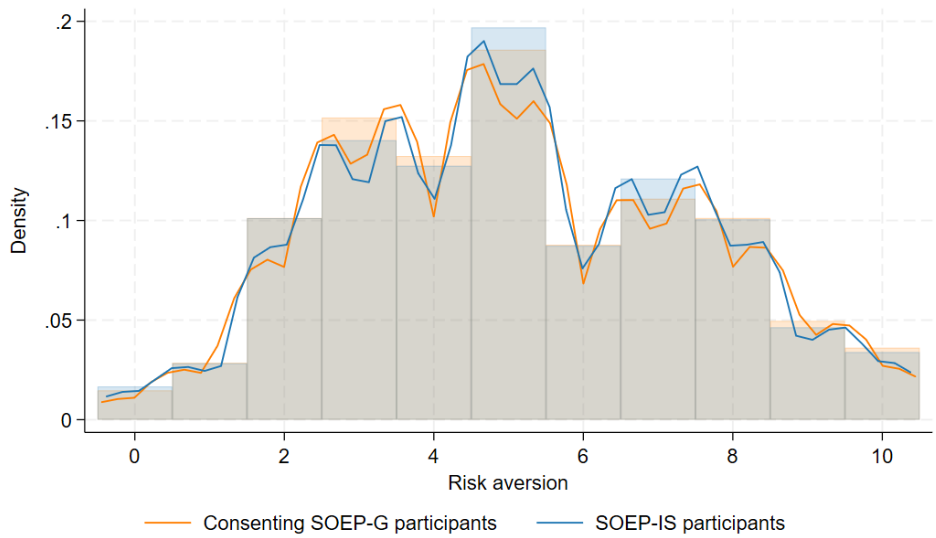
Notes: Table A20 displays the association between PGI score for cognitive skills and risk/time preference measures based on weighted OLS regressions. The weights are constructed based on the consent of SOEP-IS respondents to participate in the SOEP-G sample. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year. Robust standard errors are included in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Table A21. Weighted association between PCI CS and social preferences

	(1) Trust	(2) Pos. reciprocity	(3) Neg. reciprocity	(4) Altruism	(5) Lending help	(6) Lending belongings	(7) Lending money
PGI CS	0.094***	-0.025	-0.058***	0.017	-0.008	0.053**	0.055***
	(0.022)	(0.024)	(0.022)	(0.023)	(0.022)	(0.022)	(0.021)
Observations	2189	2197	2197	2130	2173	2166	2167
R^2	0.122	0.059	0.080	0.106	0.132	0.139	0.209

Notes: Table A21 displays the association between PGI score for cognitive skills and social preference outcomes based on weighted OLS regressions. The weights are constructed based on the consent of SOEP-IS respondents to participate in the SOEP-G sample. The regression controls for age, age-squared, parental education and occupation, the first 10 principal components of genetic data, and indicators for growing up in a two-parent household, gender and birth year. Robust standard errors are reported in parentheses. * p < 0.10, ** p < 0.05, and *** p < 0.01. *Source:* Authors' calculations based on SOEP-IS 2020.

Figure A3. Distributions of risk aversion



Notes: Figure A3 compares the distribution of risk aversion among SOEP-IS participants against that of the subset who consented to participate in SOEP-G, from which our analytical sample is drawn. The bars indicate the histogram of each sample overlying on top of one another. The lines are obtained from kernel density estimations using Epanechnikov kernel. For the SOEP-G sample coloured in orange, the bandwith used is 0.450. The bandwith in the estimation for the SOEP-IS, coloured in blue, is 0.398.

References

Angrist, Joshua D., and Jürgen S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Becker, Joel, Casper A. P. Burik, Grant Goldman, Nancy Wang, Hariharan Jayashankar, Michael Bennett, Daniel W. Belsky, Richard Karlsson Linnér, Rafael Ahlskog, Aaron Kleinman et al. 2021. "Resource Profile and User Guide of the Polygenic Index Repository." *Nature Human Behaviour* 5 (12): 1744–1758.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2013. "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences." *Journal of the European Economic Association* 11 (6): 1231–1255.

Benjamin, Daniel J., David Cesarini, Patrick Turley, and Alexander Strudwick Young. 2024. "Social-Science Genomics: Progress, Challenges, and Future Directions." NBER Working Paper 32404.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences of the United States of America* 106 (19): 7745–7750.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review* 100 (3): 1238–1260.

Dunn, Olive Jean. 1961. "Multiple Comparisons Among Means." *Journal of the American Statistical Association* 56 (293): 52–64.

Elinder, Mikael, and Oscar Erixson. 2024. "An Inquiry into the Relationship Between Intelligence and Prosocial Behaviour: Evidence from Swedish Population Registers." *Economic Journal*: 1141–1166.

Sanz-de Galdeano, Anna, and Anastasia Terskaya. 2023. "Sibling Differences in Genetic Propensity for Education: How Do Parents React?" *The Review of Economics and Statistics*: 1–44.

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge, UK: Cambridge University Press.

Koellinger, Philipp D., Aysu Okbay, Hyeokmoon Kweon, Annemarie Schweinert, Richard Karlsson Linnér, Jan Goebel, David Richter, Lisa Reiber, Bettina Maria Zweck, Daniel W. Belsky et al. 2023. "Cohort Profile: Genetic Data in the German Socio-Economic Panel Innovation Sample (SOEP-G)." *PLOS ONE* 18 (11): e0294896.

List, John A., Azeem M. Shaikh, and Yang Xu. 2019. "Multiple Hypothesis Testing in Experimental Economics." *Experimental Economics* 22 (4): 773–793.

Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2): 187–204.

Westfall, P H, and S Stanley Young. 1992. *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment*. Wiley Series in Probability and Statistics. Nashville, TN: John Wiley & Sons.

Young, Alwyn. 2018. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics* 134 (2): 557–598.