The rural-urban gap in educational inequality: Enhancement or compensation in genetic associations with educational attainment

Josué Teran Linarte, Christoph Spörlein

Article - Version of Record

Suggested Citation:
Teran Linarte, J., & Spörlein, C. (2026). The rural-urban gap in educational inequality: Enhancement or compensation in genetic associations with educational attainment. Social Science Research, 135, Article 103324. https://doi.org/10.1016/j.ssresearch.2026.103324

Social Science Research 135 (2026) 103324

Contents lists available at ScienceDirect

Social Science Research

journal homepage: www.elsevier.com/locate/ssresearch

The rural-urban gap in educational inequality: Enhancement or compensation in genetic associations with educational attainment

Josué Teran Linarte [*], Christoph Spörlein

Department of Social Sciences, Heinrich-Heine Universität Düsseldorf, Universitätsstr. 1, 40225, Düsseldorf, Germany

ARTICLE INFO

Keywords:
Rural-urban gap
Educational attainment
Gene-environment interaction (GxE)
Polygenic index (PGI)

ABSTRACT

"While educational systems have expanded, regional and social disparities in attainment remain remarkably persistent. This study integrates social stratification research with behavioural genetics to examine how macro-level regional opportunity structures and micro-level family environments moderate the genetic influence on education. Using data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) and a Polygenic Index (PGI) for educational attainment, we test a Gene-by-Environment-by-Environment (GxExE) framework. Contrary to social enhancement and compensation models, our results indicate that genetic associations with education are stable across both rural/urban contexts and varying socioeconomic backgrounds. However, we find a significant interaction between family SES and urbanicity: urban environments amplify the advantages of high-SES families, whereas rural contexts appear to act as a "leveller" that constrains opportunities regardless of social background. These findings suggest that regional macro-structures do not moderate genetic associations directly but instead condition the efficacy of social and cultural capital. Our study underscores the necessity of incorporating macro-level institutional contexts into genetically informed models of inequality."

1. Introduction

Since the 20th century, Western education systems have significantly expanded, aiming to provide universal educational attainment (Jackson, 2013). While some forms of educational inequality – such as gender disparities (Autor et al., 2016; Buchmann et al., 2008; Herd et al., 2019) – have changed dramatically during this process, other dimensions of inequality have proven more persistent (Pfeffer, 2008; Van De Werfhorst and Mijs, 2010). In particular, social and regional disparities have shown relatively little systematic change over recent decades, with children of higher socioeconomic backgrounds and urban areas consistently attaining higher levels of education in the United States (Marré, 2017). Given the long-term consequences of educational attainment over the life course, including higher income, better health and greater life satisfaction and the rural-urban gap in all these outcomes (Byun et al., 2012; Hummer and Hernandez, 2013; Kahneman and Deaton, 2010), understanding the interplay between social and regional inequality remains an important task for both sociologists and policymakers. For sociologists, the task is to uncover the specific mechanisms through which rural and urban contexts shape educational trajectories. For policymakers, this understanding is crucial to move beyond "one-size-fits-all" solutions and design targeted, place-based educational policies that can effectively address the distinct barriers faced by students in non-metropolitan areas (Showalter et al., 2019; Tieken, 2014).

Various explanations have been proposed to account for social and regional inequality in educational outcomes, including

---

[*] Corresponding author.
E-mail address: josue.teran.linarte@hhu.de (J.T. Linarte).

https://doi.org/10.1016/j.ssresearch.2026.103324
Received 5 August 2025; Received in revised form 5 January 2026; Accepted 9 February 2026

individual and contextual differences with respect to resources and aspirations (Bourdieu et al., 1977; Coleman, 1988; Sewell et al., 1969; Stocké et al., 2019). However, relatively few studies have examined social and regional inequality simultaneously, particularly in terms of how regional opportunity structures shape the mechanisms that sustain social inequality in educational attainment (Byun et al., 2012; Weβling and Bechler, 2019). Rural environments often lack essential resources, schooling options, and infrastructure – all of which have been associated with lower academic performance, graduation rates and educational attainment (Ramos et al., 2016). While these structural conditions may generally impede educational attainment in rural areas, children from higher socioeconomic backgrounds may be better positioned to compensate for these disadvantages. Families with greater resources may be able to access scarce local educational opportunities (e.g., affording specialized educational services or leveraging high-achieving social networks). As a result, rural educational constraints may amplify social inequality in educational attainment (Roscigno and Crowle, 2001).

While family characteristics are central in understanding educational inequality, sociological research has previously largely neglected that genetic differences are an important element. Like many other individual traits and behaviours, education possesses a substantial genetic component, meaning that individual differences in educational attainment can be partly attributed to genetic differences between individuals (Belsky et al., 2016; Branigan et al., 2013; De Zeeuw et al., 2015; Lee et al., 2018; Okbay et al., 2016, 2022; Polderman et al., 2015; Silventoinen et al., 2020). Indeed, the intergenerational similarity in education is partly genetically confounded, meaning that children attain similar levels of education to their parents not only because of socialization but also because they resemble each other genetically (Baier et al., 2022; Cheesman et al., 2020; Eifler and Riemann, 2022; Halpern-Manners et al., 2020; Liu, 2018; Nivard et al., 2022, 2024; Van Hootegem et al., 2025). More importantly, genes and environment work together to create differences between individuals. A growing body of research suggests that the genetic influence on educational attainment is shaped by social structures and environmental changes. For example, shifts towards more meritocratic systems and educational reforms have been linked to an increased role of genetic effects (Colodro-Conde et al., 2015; Heath et al., 1985; Pettersson, 2025; Rimfeld et al., 2018; Spörlein et al., 2025). These findings align with a gene-environment (GxE) perspective which posits that social environments can condition the extent to which genes play a role (Cheesman et al., 2025; Harerimana et al., 2025; Mills and Tropf, 2020; Silva et al., 2025; Stienstra et al., 2024; Uchikoshi and Conley, 2021).

Regional disparities may represent a similar type of GxE interaction, wherein the limited structural opportunity in rural areas constrains the extent to which genetic differences influence educational attainment. GxE also exists at the level of the family. Prior behavioural genetics studies suggest that family environments in high-status families tend to be more enriching, reducing environmental variability and thereby highlighting genetic influences on education-related outcomes (Baier and Lang, 2019; Breinholt and Conley, 2020; Ghirardi and Bernardi, 2025; Silva et al., 2025; Turkheimer et al., 2003). Conversely, among lower-status families, environmental factors play a much larger role in shaping educational outcomes suggesting that these families experience lower levels of equality of opportunity. Overall, structural conditions – whether at family or regional level – serve as vital facilitators of genetic influences, thereby contributing to the manifestation of broader patterns of inequality.

In this article, we integrate research on social and regional inequality in educational attainment with insights from the GxE behavioural genetics literature. We examine whether and to what extent an individual's social background enhances educational opportunities or compensates for their lack. Following the basic premise of equality of opportunity research using genetically informed data (Engzell and Tropf, 2019; Herd et al., 2019), we propose a gene-by-environment-by-environment interaction (GxExE), where the magnitude of genetic associations with educational attainment depends not only on the immediate family environment as captured by the social background, but also on the broader regional opportunity structure. While a growing body of research scrutinizes how environments moderate genetic influences, the environment itself is often narrowly defined as familial SES or, at most, the immediate neighbourhood (Haughbrook et al., 2016; Trejo et al., 2018). Our study expands this scope by conceptualizing the rural-urban divide as a macro-environmental context, thus bridging the gap between micro-level GxE interactions and macro-level structural conditions.

To test our hypotheses, we rely on data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) which provides detailed academic histories for a large sample of genotyped individuals. To capture genetic differences between individuals, the Add Health data provides a Polygenic Index (PGI) for educational attainment which summarises genetic associations with educational attainment for each individual (Okbay et al., 2022).

2. Literature review

A substantial and persistent urban–rural gap in U.S. educational attainment is well documented. Recent USDA reports (Farrigan et al., 2024; Pender et al., 2019) show that 35% of the urban population holds at least a bachelor's degree, compared to only 21% of the rural population (Marré, 2017). Given that rural areas comprise about one-fifth of the U.S. population – 46 million residents in 2024 (Farrigan et al., 2024) – these disparities are societally consequential. Urban areas also exhibit higher average SES compositions, indicating different opportunity structures for children across regions (Wang, 2024).

To explain these disparities, we draw on the regional opportunity structures framework (Bernard et al., 2023), which conceptualizes regional inequalities as arising from the interaction of individual resources with the availability, accessibility, and quality of context-specific opportunities. We apply this perspective to reconstruct arguments centred on resources and aspirations.

2.1. Resources and aspirations

Families differ systematically in their economic, cultural, and social resources (Boudon, 1974; Jackson, 2013; Stocké et al., 2019). Higher-SES families provide educational support – private tuition, extracurricular activities, enriched learning environments – that foster academic achievement (Baier and Lang, 2019; Bourdieu, 1986; Lareau, 2011; OECD, 2023; Sirin, 2005). Resource-poor families face constraints that generate home environments less aligned with school demands, potentially hindering development (Cheadle and Amato, 2011; Hoff, 2006; Uchikoshi and Conley, 2021).

These mechanisms matter strongly in regional contexts, where educational decisions are shaped by local constraints (Jackson, 2013). Rural areas often provide fewer curricular options and weaker labour markets (Byun et al., 2012; Roscigno and Crowle, 2001), amplifying the negative consequences of low family SES. Empirical work shows that lower SES strongly predicts rural students' reduced postsecondary outcomes and higher dropout rates relative to urban peers (Adelman, 2006; Paasch and Swaim, 1995).

Resource disparities are also evident at the school level (Apple et al., 2010; Banerjee et al., 2007; Behrman et al., 2010; Behrman and Birdsall, 1983; Naylor and Smith, 2004; Fertig and Wright, 2005). While evidence on rural-urban school differences is mixed (Blackwell and McLaughlin, 1998; Howley and Gunn, 2003; Sewell et al., 1969; Williams, 2005; Young, 1998), rural students disproportionately come from families with lower parental education and fewer financial resources (Farrigan et al., 2024). Because U. S. public schools rely heavily on local property taxes, more affluent communities generate greater funding and higher-quality facilities, while rural schools generally receive fewer resources per pupil (Chetty and Friedman, 2010; Showalter et al., 2019). Consequently, rural schools tend to be smaller, less well-equipped, and offer fewer academic programs (Brown and Schafft, 2018; Tayyaba, 2012).

Aspirations constitute another key mechanism (Sewell et al., 1969; Stocké et al., 2019). Higher-status families cultivate high educational expectations as part of maintaining social position, both directly and through exposure to information, role models, and cultural contexts that reinforce educational success (Chohan and Khan, 2010; Eccles and Wigfield, 2002; Hartung and Hillmert, 2019; Lareau, 2011). However, aspiration formation varies regionally. Rural areas – marked by economic disadvantage and limited structural opportunities – tend to foster lower educational aspirations, even among relatively advantaged families (Koricich et al., 2018). Mismatches between local labour-market skill demands and higher-education qualifications further reduce perceived returns to college (Brown and Schafft, 2018). Empirical evidence shows that rural students remain less likely than urban peers to attend college (McCracken and Barcinas, 1991). In contrast, urban contexts provide richer informational environments and more frequent contact with highly educated individuals, thereby supporting higher aspirations (Barone et al., 2016; Hartung and Hillmert, 2019).

2.2. *Gene-environment interactions: enhancement and compensation*

The effectiveness of resources and aspirations depends partly on individuals' general abilities, which have a substantial genetic component (Ayorech et al., 2019; Krapohl et al., 2014; Mills and Tropf, 2020; Plomin and Von Stumm, 2018; Polderman et al., 2015; Savage et al., 2018). Twin studies attribute roughly half of the variance in cognitive and educational outcomes to genetic differences (De Zeeuw et al., 2015; Polderman et al., 2015). Similarly, studies using Polygenic Index (PGI) underscore strong positive associations, with a one standard deviation PGI increase predicting up to one additional year of schooling (Domingue et al., 2015; Ghirardi et al., 2024; Lin, 2020; Okbay et al., 2022; Procopio et al., 2025). Yet, these associations are not immutable; rather, they vary depending on social conditions – such as educational expansion, political change, or normative shifts – constituting gene–environment interactions (Cheesman et al., 2022; Erola et al., 2020; Fraemke et al., 2025; Ghirardi and Bernardi, 2025; Herd et al., 2019; Knigge et al., 2022; Lahtinen et al., 2023; Ruks, 2022; Stienstra et al., 2024; Stienstra and Karlson, 2023; Trejo et al., 2018). From a theoretical perspective,

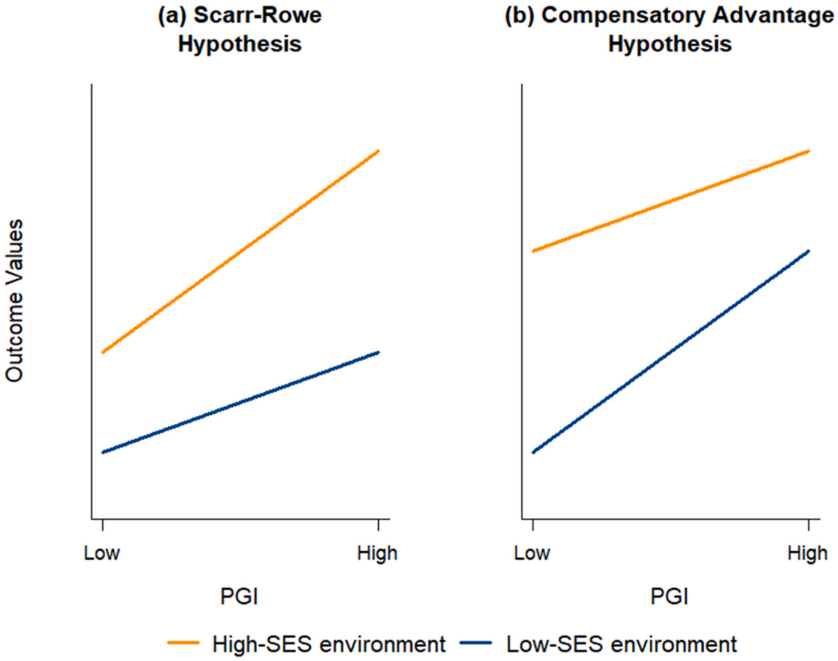


Fig. 1. Two hypotheses on the interaction between PGI for education and family SES on educational outcomes.

the magnitude of genetic influence serves as a key indicator of equality of opportunity, as stronger genetic effects imply weaker associations with ascribed characteristics such as SES (Baier, Eilertsen, et al., 2022; Baier et al., 2022; Engzell and Tropf, 2019; Harden, 2021; Plomin, 2019; Rowe et al., 1999; Scarr-Salapatek, 1971). Importantly, genetic propensity does not determine achievement but interacts dynamically with contextual conditions (Morris et al., 2020).

Two competing GxE models are central: *enhancement* and *compensation*. The enhancement model (Rowe et al., 1999; Scarr-Salapatek, 1971) posits that advantaged environments strengthen genetic associations, whereas resource-poor environments suppress them (Giangrande et al., 2019; Tucker-Drob and Bates, 2016; Turkheimer et al., 2003; Woodley of Menie et al., 2018). Empirical support for this model appears more consistent in the U.S. compared to Europe or Australia (Ghirardi et al., 2024). Conversely, the compensatory model argues that high-SES families offset lower genetic abilities through intensive investments to prevent downward mobility (Arold et al., 2022; Bernardi, 2014; Bernardi and Triventi, 2020; Dierker and Diewald, 2025). These mechanisms yield distinct empirical signatures: enhancement anticipates widening SES gaps at higher PGI levels, whereas compensation anticipates widening gaps at lower PGI levels (see Fig. 1).

Thus far, neither hypothesis has been directly tested in a rural-urban setting, where regional opportunity structures simultaneously shape family SES effects and the strength of genetic associations. Rural regions exhibit fewer educational resources and lower aspirational climates, raising the question of how families of different backgrounds navigate these constraints.

Following enhancement framework, providing enriched learning environments is significantly more difficult in rural areas, where schools have fewer resources, extracurriculars are limited, and community-level aspirational cues are weaker. Consequently, high-SES rural families face greater barriers to cultivating children's potential than their urban counterparts. We thus expect the SES-stratified PGI-attainment association to display a stronger enhancement pattern in urban areas, producing larger SES gaps at higher PGI levels.

In contrast, the compensatory model predicts that high-SES families invest to offset low abilities. Because remedial services are scarcer in rural settings, such compensatory investments are more challenging overall. However, high-SES families should still be better positioned than others to access scarce resources. Accordingly, we expect compensation dynamics to appear more strongly in rural areas, generating larger SES gaps at lower PGI levels. These hypotheses do not imply that either pattern is exclusive to one context; rather, they suggest that regional structures make one investment strategy more feasible than the other.

3. Data & methods

3.1. Data

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal and nationally representative study of adolescents in the United States who were in grades 7-12 during the 1994-1995 school year. The initial sample was drawn from a probability sample of 80 high schools and 52 middle schools, stratified by region, urbanicity, school type, ethnic mix, and size. Data were collected across five waves: Wave 1 1994-1995 (N = 20745), Wave 2 1996 (N = 14738), Wave 3 2001-2002 (N = 15179), Wave 4 2008-2009 (N = 15701), and the most recent Wave 5 2016-2018 (N = 12300). The surveys include in-depth interviews with participants, as well as questionnaires administered to parents and school representatives. The study also includes an oversample of siblings, who were genotyped at Wave 4 using saliva samples to obtain genetic markers.

3.2. Sample

The final analytical sample was constructed through the following selection process. Consistent with standard practice in genomic research, we restricted the sample to individuals of European genetic ancestry, as Polygenic Indices are predominantly derived from Genome-Wide Association Studies (GWAS) conducted on European-ancestry populations. From this subset, individuals for whom PGIs and survey design data were available were selected. The proportion of missing values in the sample necessitated specific analytical strategies (Fig. 1A in Appendix A). Employing a listwise deletion approach would have entailed the elimination of a substantial number of individuals and statistical power. Therefore, to mitigate bias of a complete-case analysis approach and leverage the available data more efficiently, we employed a multiple imputation procedure to handle missing values. The method produces unbiased estimates of the true standard errors and adjusts the variance of an estimator based on a sampling function, such that the uncertainty created by the missing information is accounted for correctly (Van Buuren, 2018). To maximize efficiency and the computational power, 25 imputed datasets were created with 20 iterations using the *mice* package in R. Variables related to genetic measures and survey design were not imputed. The pre- and post-imputation diagnostics are available in Appendix A. To test if the imputation biased our results, we also ran the same analysis using listwise deletion, which results can be found in the sensitivity analysis.

Following imputation, we randomly selected one individual per family within each of the 25 imputed datasets to establish a between-family design and maximize statistical power for detecting the gene-environment interaction. We are aware that a between-family design can produce biased estimates due to gene-environment correlation (rGE); however, the data are not suitable to run a more rigorous within-family design. We discuss the potential limits arising from this decision in the conclusions.

As Abdellaoui et al. (2019) pointed out, individuals with higher PGI levels for educational attainment tend to migrate out of economically deprived regions. In the U.S. context, this dynamic is exacerbated by the geographic distribution of educational

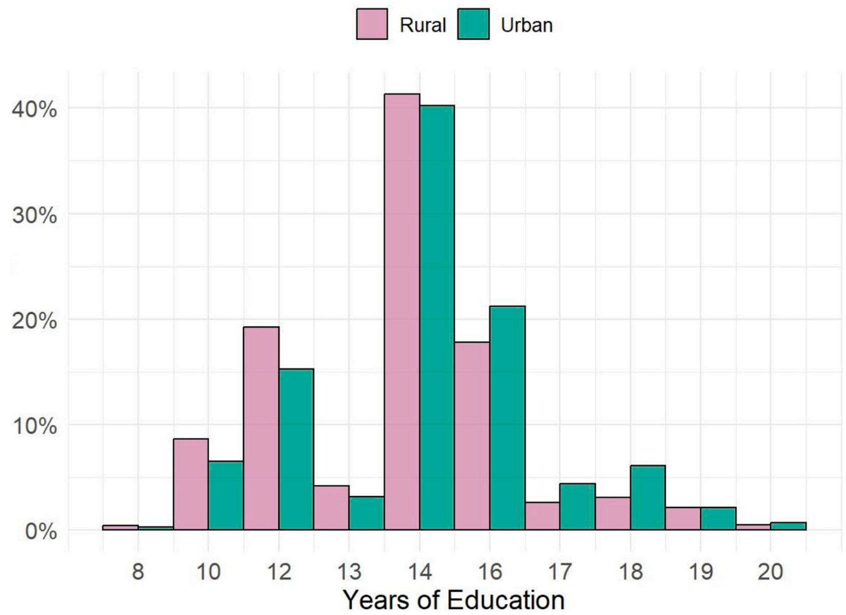


Fig. 2. Histogram of educational attainment by region.

infrastructure (Hillman et al., 2021). While we acknowledge that excluding regional migrants may induce selection bias, including them would introduce significant measurement error regarding the environmental exposure (Urban vs. Rural), obscuring the specific GxE mechanism we aim to test. Consequently, we restricted the sample to individuals who remained in their region until high school completion. As a result, 120 individuals classified as "movers" were dropped (for classification details, see Appendix B), representing just 2% of the sample, which suggests that the potential for selection bias is minimal. After applying these restrictions, we ran the analysis on a final analytical sample of N = 4915.

3.3. *Measures*

3.3.1. *Dependent variable*

Educational Attainment. We operationalize educational attainment as the total number of years of education completed at Wave 4, when respondents were, on average, over 27 years old and most had completed their formal education. The variable was derived from the survey question, "What is the highest level of education that you have achieved to date?". Following the methodology of Domingue et al. (2015) we converted these categorical responses into a continuous measure of years of schooling. In this sample, the average educational attainment is 14.2 years.

3.3.2. *Independent and moderating variables*

Polygenic Index for Educational Attainment (PGI-EA). We use the PGI for educational attainment provided by Add Health at Wave 4. A Polygenic Index is a summary score that aggregates the thousands of small associations of genetic variants with the outcome. The specific variants correlating with educational attainment and their effect sizes were estimated in independent genome-wide association studies (GWAS). As such, a PGI represents the estimated additive genetic association for a phenotype (Okbay et al., 2022). The PGI for Educational Attainment (PGI-EA) used in this study was derived from the GWAS by Lee et al. (2018).[1] GWAS and the resulting PGI cover by no means all genetic associations with an outcome as the focus on common genetic markers missing out on rare variants (Young, 2022). Moreover, PGIs have been shown to also capture social or environmental variation thereby showing some degree of social confounding (Burt, 2024). For these reasons, PGIs should be considered imperfect approximations of genetic associations. For the purpose of our analyses, the PGI-EA was standardized to a mean of 0 and a standard deviation of 1.

Family Socioeconomic Status (SES). Family SES is operationalized using a constructed variable by Add Health at Wave 1, based on parent-reported information on four correlated indicators: parental education, parental occupation, household income, and household receipt of public assistance. Add Health employed a principal components analysis to create a continuous factor score. We

---

[1] While a newer GWAS summary statistic is available (Okbay et al., 2022), this score has not yet been integrated into the public Add Health release. Consequently, we rely on the Lee et al. (2018) score, consistent with other recent applications of this dataset (e.g., Weng, 2025). Furthermore, we restrict our analysis of educational attainment to Wave 4 (ages 28–30) rather than including Wave 5. This decision minimizes panel attrition while ensuring that the window of observation is sufficient to capture completed formal education for the vast majority of the cohort, consistent with prior literature (Domingue et al., 2015).

standardized it to a mean of 0 and a standard deviation of 1.

Urban vs. Rural Residence. To capture the regional context, we relied on Rural-Urban Commuting Area (RUCA) codes derived from the 1990 U.S. Census and linked to respondents' addresses at Wave I. RUCA codes classify census tracts using measures of population density, urbanization, and daily commuting flows. While the original data provides a 10-point scale of regional differentiation (U.S. Department of Agriculture, Economic Research Service, 2025a), we followed the Add Health guidelines to aggregate these into four primary categories (Godwin, 2024): (1) Urban, N = 3763; (2) Large Rural City/Town, N = 333; (3) Small Rural Town, N = 441; and (4) Isolated Small Rural Town, N = 378. Due to limited sample sizes in the rural sub-categories, we further operationalized the data into a binary variable. The last three categories were combined into a single "Rural" indicator, while the first remained "Urban." This binary classification is consistent with approaches used in prior literature (e.g., Lawrence et al., 2017), and also part of the census guidelines (U.S. Department of Agriculture, Economic Research Service, 2025b). To ensure this aggregation process, we also conducted a sensitivity analysis using the original four-category classification. In our analytical sample, 3762 individuals (77%) resided in urban areas, while 1153 (23%) resided in rural areas. Notably, this distribution aligns closely with U.S. Census Bureau data from the 1990s (U.S. Bureau of the Census, 1992).

3.3.3. *Control variables*

All models include a set of standard control variables. We control for the first 10 genetic principal components (PCs) to account for population stratification, a standard procedure in genetic association studies (Abdellaoui and Verweij, 2021; Price et al., 2006). We also control for biological sex and age.

Descriptive statistics for the analytical sample can be found below in Table 1.

3.4. *Analytical strategy*

To test our hypotheses, we use Ordinary Least Squares (OLS) regression to estimate a three-way gene-by-environment-by-environment (GxExE) interaction model.[2] Our analytical strategy proceeds in three steps. First, we estimate a baseline linear regression model for educational attainment that includes the PGI-EA, family SES, and the urban/rural indicator, alongside all control variables (vector Xi) (Eq. (1)).

$$\mathrm{Yi} = \beta 0 + \beta 1\mathrm{PGI} - \mathrm{EAi} + \beta 2\mathrm{SESi} + \beta 3\mathrm{Urbani} + \beta 4\mathrm{Xi} + \epsilon \mathrm{i} \tag{Eq.1}$$

Second, we introduce the GxExE interaction term between the PGI-EA, family SES, and the urban/rural indicator (Eq. (2)). To control for potential confounders, we also include interaction terms between our main variables of PGI, SES, Urban and all control variables (Keller, 2014).

$$\begin{aligned}\mathrm{Yi} = {} & \beta 0 + \beta 1\mathrm{PGIi} + \beta 2\mathrm{SESi} + \beta 3\mathrm{Urbani} + \beta 4(\mathrm{PGIi} \times \mathrm{SESi} \times \mathrm{Urbani}) + \beta 5\mathrm{Xi} + \beta 6(\mathrm{PGIi} \times \mathrm{Xi}) + \beta 7(\mathrm{SESi} \times \mathrm{Xi}) + \beta 8(\mathrm{Urbani} \times \mathrm{Xi}) \\ & + \epsilon \mathrm{i}\end{aligned} \tag{Eq.2}$$

We estimate all models with robust standard errors clustered at the school level, stratified by region, and weighted, as suggested by the Add Health documentation (Chen, 2014).

All analyses were conducted on the 25 imputed datasets and the results were pooled using Rubin's rules (Van Buuren, 2018), which produce the final standard errors by combining the average within-imputation variance with the variance between imputed estimates, thereby explicitly accounting for the uncertainty introduced by the imputation process (Rubin, 1987). For each of the models tested, we extracted the coefficients of interest along with their corresponding standard errors. These values were then combined to produce a single pooled estimate, a pooled standard error, and a 95% confidence interval. Furthermore, the R^2 values were pooled using Fisher's r-to-z transformation (Harel, 2009). To ensure reproducibility, the R script for all analyses presented in this paper is publicly available at https://github.com/Josuelinter/GxE-Rural-Urban.

4. Results

We begin with a descriptive analysis, examining the distribution of educational attainment, family SES, and the PGI-EA across urban and rural contexts.

As shown in Fig. 2, the difference in the average years of education between the two areas is minimal. However, an analysis of the distribution reveals subtle but substantive distinctions. Rural areas exhibit a higher concentration of individuals with lower educational attainment (10–12 years), whereas urban areas display a thicker tail at the upper end of the distribution (14+ years), indicating a higher prevalence of post-secondary and postgraduate education. It is important to note that datasets with genetic data are often subject to selection bias, over-representing individuals with higher educational attainment. The lack of a strong systematic difference

---

[2] Given the sample size requirements for detecting complex interactions, we conducted a post-hoc sensitivity power analysis following the approach of Ghirardi et al. (2024). We calculated the minimum incremental R-squared required to yield a statistically significant result in an F-test, assuming a statistical power of 0.80 and an alpha of 0.05. With our final analytical sample (N = 4915), the study is sufficiently powered to detect an incremental increase in R-squared as small as 0.0012 (0.12%) for the interaction terms.

Table 1
Descriptive statistics of the analytical sample.

Variables	Mean	SD	Min	Max
Educational Attainment	14.2	2.16	8	20
PGI-EA	0	1	−4.14	3.40
Family SES	0	1	−3.96	2.77
Urban	0.77		0	1
Female	0.53		0	1
Age	28.4	1.7	24	34
N	4915			

Notes: Mean, standard deviations, min and max values. Descriptives of the first 10 principal components of genetic ancestry not shown.

in our sample could, therefore, be part of this mechanism. The distribution of family SES reveals a clearer pattern (see Fig. 3). Lower-SES families are more prevalent in rural areas, while the urban sample comprises a significantly larger proportion of families from higher socioeconomic groups. Fig. 4 displays the distribution of the PGI-EA showing essentially no difference across the contexts.

To formally test these distributions, we employed both *t*-test and Kolmogorov-Smirnov (K–S) tests. The t-tests indicated a statistically significant difference in means (p-value <0.05). Note that such results are likely driven by the large sample size. However, the K–S test for the PGI-EA was not statistically significant, indicating that the overall shape of the genetic distribution is identical across rural and urban populations. This is an important finding as it will enable us to highlight the influence of social environmental differences in the absence of meaningful genetic differences across rural and urban areas.

Before turning to the regression models, we addressed the potential for Gene-Environment Correlation. Consistent with previous literature (D. Belsky et al., 2018; Biroli et al., 2025; Zhou et al., 2024), we found that family SES is not independent of an individual's genotype (see Fig. 5). The bivariate analysis indicates a positive correlation between PGI-EA and family SES ($r = 0.24$). Given this substantial overlap – and the fact that family SES is a strong predictor of education ($r = 0.46$) – we interpret our results recognizing that genetic associations may partly reflect environmental advantages, and conversely, environmental effects may partially capture genetic confounding. The consequences of such correlation will be discussed in the limitations.

Table 2 presents the results of our regression models. In the baseline model (Model 1) a one standard deviation increase in the PGI-EA is associated with an increase of 0.61 years of education. When we introduce the family SES indicator, the PGI association decreases (Model 2: $\beta = 0.38$ SD), as family SES is also positively associated with educational attainment. Specifically, one standard deviation in family SES is associated with 0.89 years more of schooling. This attenuation reflects the rGE discussed above since a portion of the PGI effect in Model 1 is mediated through, or confounded by, family socioeconomic background. Model 3 allows the association between the PGI to vary across family SES but does not show any statistically significant moderation.

We then test our main hypothesis on the GxExE interaction. Model 4 represents the association between the PGI-EA and educational attainment conditioned on family SES and regional setting. Contrary to our hypothesis that investigates the different pattern of *enhancement* or *compensation* hypothesis, we did not find a statistically significant result ($\beta = 0.048$ SD, p-value $= 0.577$). This indicates

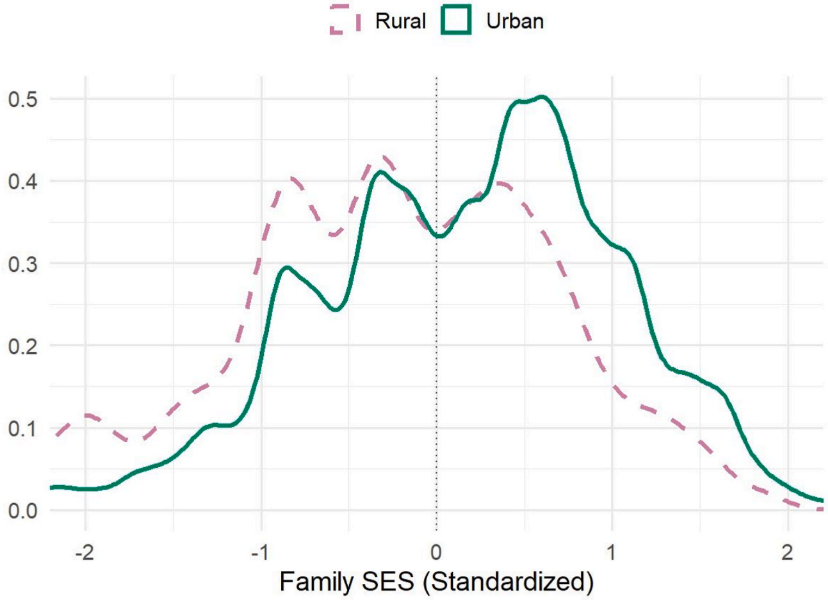


Fig. 3. Average distribution of Family SES by region.
Notes: Family SES standardized with mean 0 and SD 1.

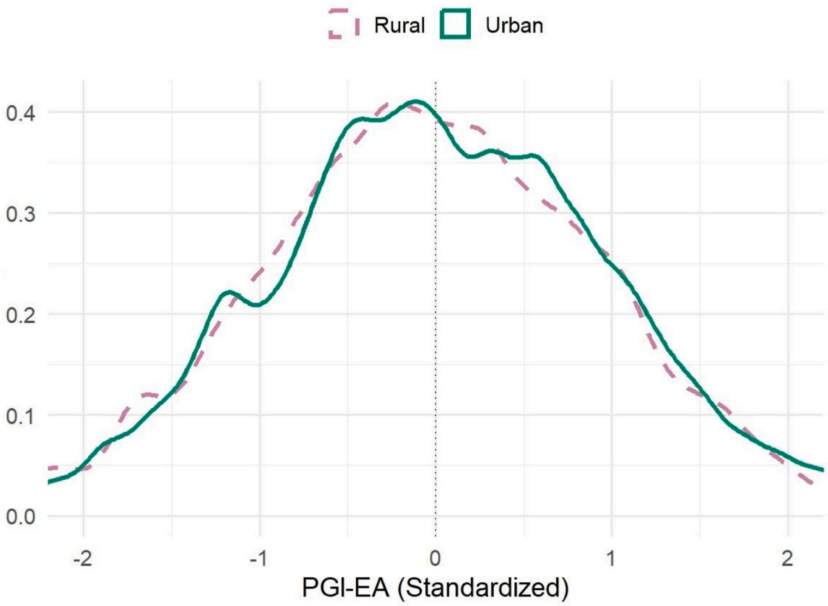


Fig. 4. Average distribution of PGI-EA by region.
Notes: PGI-EA standardized with mean 0 and SD 1.

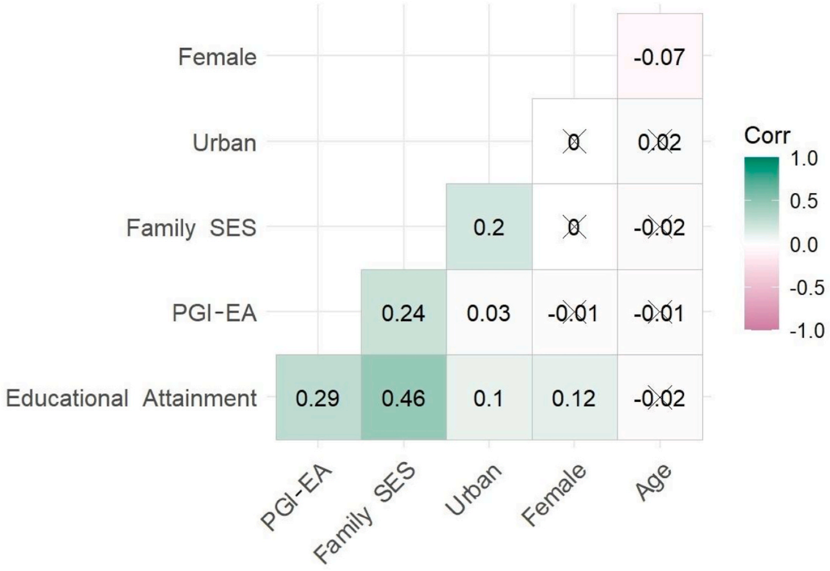


Fig. 5. Correlation matrix.
Notes: Crossed cells ("X") indicate non-significant coefficients (p-value >0.05).

that the moderating role of family SES on the genetic association with education does not differ significantly between rural and urban contexts. However, while the genetic interaction was null, Model 4 reveals a significant interaction between family SES and Urban indicator ($\beta = 0.204$ SD, p-value <0.05). This finding suggests that the relationship between family background and educational attainment is conditional on the living area.

Fig. 6 illustrates the nature of the three-way interaction by plotting the predicted educational attainment across different levels of family SES and the PGI. The findings highlight two key patterns. In rural areas (left side), the slopes remain parallel across the SES spectrum, indicating that high-SES families in rural areas do not gain any disproportionate advantage in converting genetic potential into attainment compared to lower-SES families. In urban areas (right side) we observe a steeper gradient for family SES. While the genetic slopes do not significantly fan out, confirming the null three-way interaction, the gap between low and high SES is significantly wider in urban settings. Consequently, the data seem to support that urban environments facilitate high-SES families. A result that

Table 2
Regression models of the gene x family SES × urban interaction on educational attainment (N = 4915).

Educational Attainment	(1)	(2)	(3)	(4)
PGI-EA	0.617***	0.383***	0.343***	0.447**
	(0.044)	(0.035)	(0.055)	(0.072)
Family SES		0.898***	0.917***	0.757***
		(0.052)	(0.068)	(0.084)
Urban (ref: No) Yes		−0.152	−0.152	0.020
		(0.124)	(0.122)	(0.152)
PGI-EA × Family SES			0.071	0.042
			(0.039)	(0.072)
PGI-EA × Urban				−0.138
				(0.072)
Family SES × Urban				0.204*
				(0.095)
PGI-EA × Family SES × Urban				0.048
				(0.086)
Control variables		X	X	X
Control interaction			X	X
Constant	14.145***	14.038***	14.013***	13.845***
	(0.087)	(0.112)	(0.114)	(0.129)
R-squared	0.062	0.255	0.261	0.271

Notes: Controlled for the PCs, sex, and age in all models. Further interactions controlled in model 4. Robust standard errors (SE) in parentheses; SE clustered by school classes; ***P < 0.001, **P < 0.01, *P < 0.05. All continuous independent variables are z-standardized prior to the analyses. Multiple imputation used to fill missing data (m = 25).

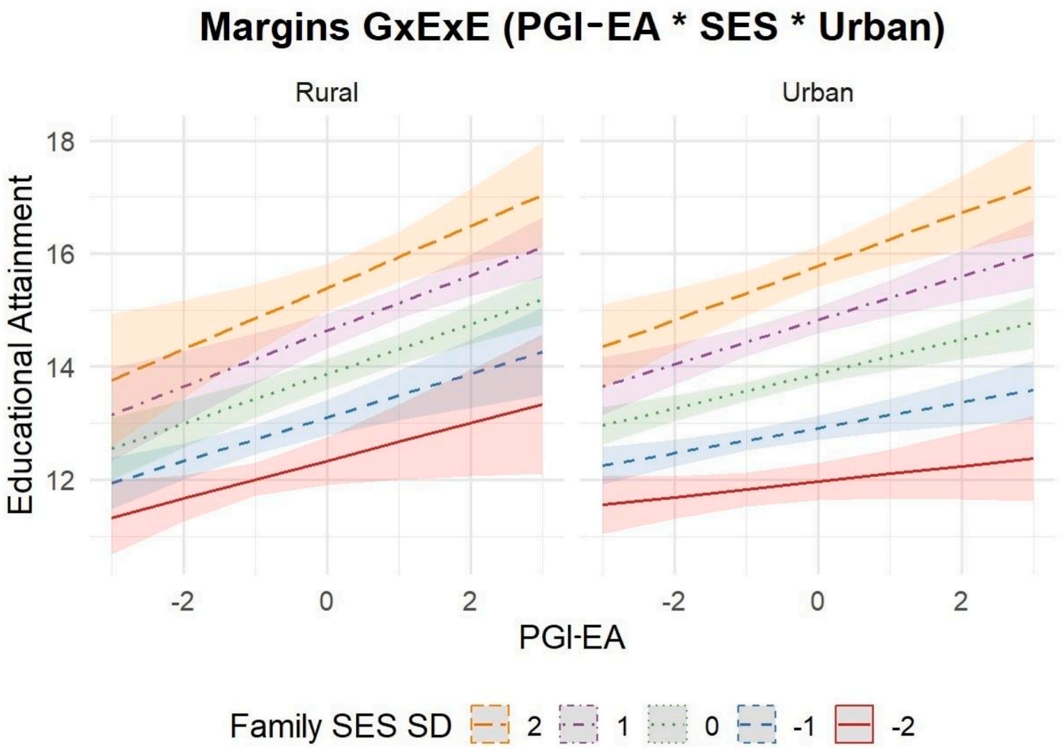


Fig. 6. Socially stratified relationship between PGI and educational attainment across rural and urban areas.
Notes: Controlled for the PCs, sex, age and interaction with Keller adjustments. Numeric values of family SES indicate standard deviation changes.

underlines how such families may be more effective at leveraging their resources to increase educational attainment, independent of their children's genetic score.

4.1. Sensitivity analyses

We replicate the analysis using a continuous measure of neighbourhood status. This approach allows us to test that the observed differences between rural and urban environments reflect structural constraints rather than simple within wealth variation. We replaced the binary Urban-Rural indicator with a continuous measure of neighbourhood income. Table 1B (in Appendix B) shows the main results. The new model yields two critical insights. First, we observe a significant main effect of neighbourhood income (β = 0.21, p-value = 0.010), meaning that residing in a richer neighbourhood is independently related with educational attainment regardless of genetic associations or family background. Second, the three-way interaction between PGI-EA, family SES, and Neighbourhood Income is not statistically significant (p-value = 0.907), just as the PGI-EA x Neighbourhood. This contrasts with our primary model, where the interaction with the Urban context was significant. Such results suggest that the main analysis is robust, and that the moderation of this effect is specific to urban dimensions. These findings indicate that the specific stratification patterns observed in our main analysis are not artifacts of neighbourhood wealth. Instead, the urban effect likely captures distinct structural characteristics that exist in urban cities but are absent in rural areas, even those of comparable economic standing.

Second, we performed a robustness check using the full four-category RUCA classification (Table 2B in Appendix B): Urban (Reference), Large Rural City, Small Rural Town, and Isolated Rural Town. The analysis revealed that the main effects of the polygenic score and SES remained consistent with the main model. We observed a significant negative interaction between SES and the Isolated Rural Town category (β = −0.399, p-value <0.001). However, no significant interactions were found regarding the polygenic score across the varying rural categories, nor were any three-way interactions statistically significant, which is in line with our main analysis. It is important to note that the sample sizes for the disaggregated rural categories are small compared to the Urban category (76% of sample). The resulting scarcity of data in the rural subgroups likely reduced statistical power and, indeed, produced large confidence intervals, particularly for higher-order interactions.

And finally, to assess the robustness of the main results, we replicated the final model using listwise deletion and compared it to the multiple imputation estimates (Table 3B in Appendix B). The listwise deletion approach resulted in a substantial reduction in sample size, dropping from N = 4915 in the imputed model to N = 3573, a loss of approximately 27% of the observations. As shown in Table 3B, the direction and magnitude of the key interaction term (SES x Urban) remain consistent across both approaches (β IM = 0.20; β LS = 0.17). However, the effect in the listwise model fails to reach statistical significance (p-value = 0.07). This discrepancy is attributable to the reduced statistical power inherent in the smaller listwise sample. Consequently, the MI model is preferred as it maximizes the use of available information, preserves statistical power, and corrects for potential biases associated with discarding incomplete cases.

5. Summary and conclusion

This article integrates G × E research with the study of regional disparities to move beyond analyses limited to the family environment. We seek to disentangle the relationship between macro-level contexts such as urban and rural areas and micro-level processes. Our findings using Add Health data confirm that the PGI for educational attainment and family socioeconomic status are robust, predictors of education. However, the place of residence was not a significant predictor on its own. Critically, our analysis reveals that the association between genetic influences and educational attainment is not moderated by family resources or regional settings. We found no support for a G × E interaction in this context; consequently, our data do not support either the Scarr-Rowe (enhancement) or compensation models of interaction. This aligns with recent work running similar analyses and data (e.g., Ghirardi and Bernardi, 2025; Trejo et al., 2018).

What appears to matter more in this scenario is the interplay between family SES and the regional context. We observe a larger gap between social classes in urban areas compared to rural ones, independent of the PGI indicator. Hence, urban environments are resource-rich enough to allow high-SES families to leverage their advantages effectively, whereas rural environments may act as a leveller by constraining opportunities for all.

This dynamic could stem from structural constraints; rural areas often lack the educational resources and higher education institutions found in cities, requiring individuals to migrate to achieve advanced degrees. While there is no direct association between the urban-rural indicator alone and attainment, the significant relationship between family SES and urban residence implies that presence in an urban area is necessary but insufficient. To capitalize on urban opportunities, families must possess the right resources – cultural and social capital – potentially facilitating access to better schools or fostering higher educational aspirations due to proximity to universities (Wang, 2024; Zahl-Thanem and Rye, 2024).

In rural contexts, however, individuals may have to rely more heavily on their own specific combination of socioeconomic and genetic resilience to navigate the life course. While not statistically confirmed here, this pattern would align with research on deprived neighbourhoods (Harerimana et al., 2025; Hartung and Hillmert, 2019; Silva et al., 2025), suggesting that high PGI might be associated with resilience traits that help individuals mitigate the lack of regional opportunities.

Ultimately, our results underscore the profound importance of the institutional context. The patterns observed in the U.S. likely differ from those in countries with more egalitarian educational systems (Cheesman et al., 2022, 2025). The highly stratified nature of American society creates a context that enables higher-SES families to translate their resources into educational advantages more directly.

The following limitations should be considered in future research. First, the Polygenic Index is an imperfect measure of biological markers (Burt, 2024). Although PGIs reflect a linear predictor of an outcome, they are subject to biases from population stratification, assortative mating, and genetic nurture (Alemu et al., 2025; Blanc et al., 2025). Consequently, they depend on the environmental factors and demographics of the discovery sample, meaning the association between a PGI and an outcome cannot be interpreted as an immutable biological relationship. While we addressed population stratification by controlling for genetic ancestry following standard practices (Abdellaoui and Verweij, 2021), bias remains a concern. Future research should prioritize within-family or trio designs when data allow, or include family fixed effects in GWAS analyses to reduce bias from assortative mating and gene-environment correlations (Biroli et al., 2025).

Second, gene-environment correlation remains a significant challenge. The gold standard in behavioural genetics is to control for familial confounders to isolate causal effects, but the Add Health dataset lacks the parental genotype data necessary for a within-family design. Consequently, we relied on a between-family design, meaning our results may be biased by passive rGE. For instance, parents with genotypes conducive to education may provide environments that foster learning (genetic nurture) and are also more likely to have higher SES and reside in specific regions (Breinholt and Conley, 2020; Kong et al., 2018). Since SES acts as a mediating mechanism in many G × E studies and influences residential selection, it is plausible that the relationship between family SES and urban residence is partly driven by genetic factors rather than purely social ones. We therefore caution that our results should be interpreted as associations rather than causal relationships.

Furthermore, active rGE may also be at play. Individuals with a higher genetic propensity for education may selectively migrate to urban areas to access higher education, creating a correlation between genetics and regional environment (Abdellaoui et al., 2019, 2022). Although we restricted our sample to individuals who remained in the same environment until age 18 to mitigate this, avoiding selection bias entirely is difficult.

Third, the issue of endogeneity persists. G × E interactions are most clearly identified when genetic endowments and environmental factors are exogenous to each other. However, in the current state of the art, exogenous genetic measures are rare (Biroli et al., 2025), and our environmental measures (SES and residence) are likely endogenous, potentially reflecting parental or individual genetics. Future research should aim to exploit exogenous environmental shocks, such as social reforms, to better identify these interactions and distinguish true G × E from G × G or E × E effects.

Fourth, the Add Health data are left-censored, as the study began when participants were already adolescents. We lack data on early childhood environments, a critical period for cognitive development when environmental factors have a fundamental impact (Ruiz et al., 2016). A more complete geographical history would enhance accuracy. Additionally, our analysis relies on a broad governmental definition of urban and rural areas. Some areas classified as rural may exhibit urban characteristics regarding opportunities, and vice versa. Future work would benefit from multidimensional measures of regional context.

Finally, sample size constraints and the age of the data must be noted. A common limitation in gene-environment interaction research is the requirement for very large sample sizes to detect complex effects (Domingue et al., 2020). To address this, we conducted a post-hoc sensitivity power analysis. The analysis indicated that our study was sufficiently powered to detect very small effects. However, given the complexity of GxExE models, the risk of spurious findings or measurement error remains a limitation, highlighting the importance of replication with larger samples. Moreover, the regional contexts from the initial waves are now dated; contemporary policies addressing “left-behind” regions (Rodríguez-Pose, 2018) may have altered the dynamics observed in our study.

Notwithstanding these limitations, we contribute to the social stratification literature by demonstrating that the interplay between genes and environment is contingent on macro-level structures. We highlight the necessity of not only integrating genetic measures into stratification models but also of using theories from social stratification to disentangle the social mechanism behind behavioural genetic findings. We believe that a greater integration of sociological theory is crucial for a clearer understanding of how life course is shaped by both genes and social environments.

CRediT authorship contribution statement

Josué Teran Linarte: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Christoph Spörlein:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Funding sources

This work was supported by the Deutsche Forschungsgemeinschaft (grant number SP 1569/5-1)

Appendix A

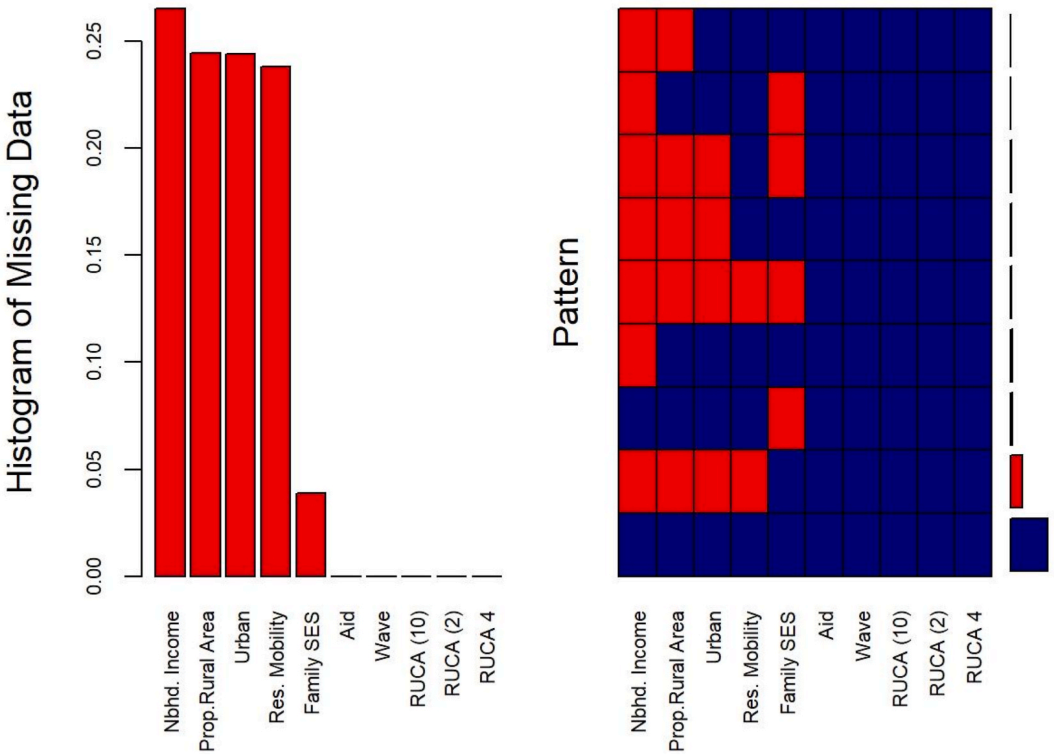


Fig. 1A. Diagnosis pre-imputation.

Notes: The bar chart (left) indicates the proportion of missingness per variable, and the grid (right) depicts the combination of missing values between variables. Red indicates missing values. The counts of missing observations are as follows: Neighbourhood Income (n = 1444), Proportion Rural (n = 1330), Residential Mobility (n = 1296), and Family SES (n = 211). The Urban variable (n = 1328) shown here is a pre-registered measure utilized solely as an auxiliary variable to inform the imputation process; it is distinct from the final urban/rural classification used in the main analysis.

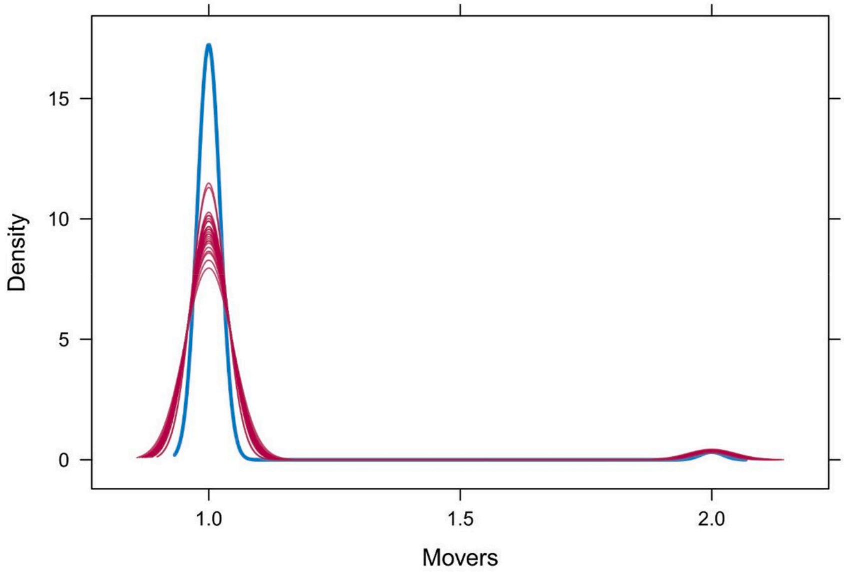


Fig. 2A. Diagnostic plot after-imputation.

Notes: Differences in the distribution of imputed data and the observed data for the movement variable. It is the variable with most missing values in the main analysis. Imputation method *logreg*.

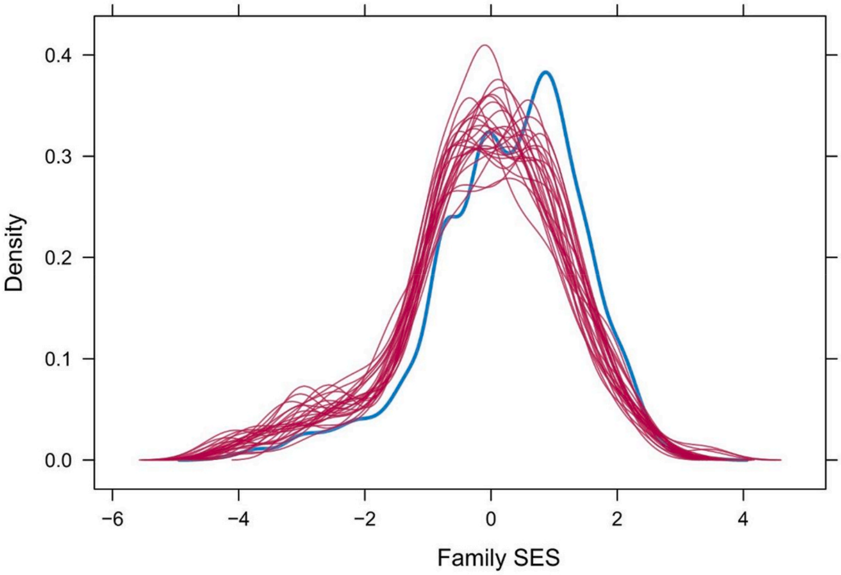


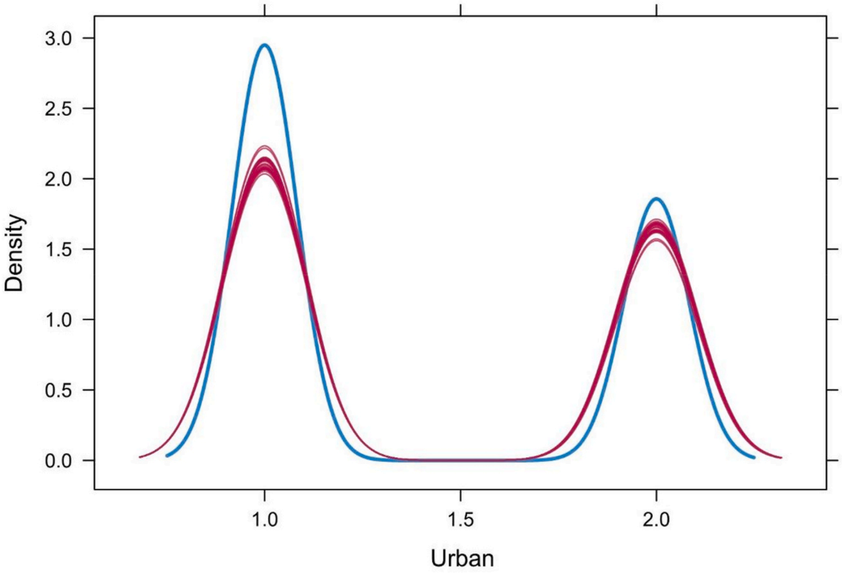
Fig. 3A. Diagnostic plot after-imputation.

Notes: Differences in the distribution of imputed data and the observed data for the family SES variable. Imputation method *pmm.*

Fig. 4A

Diagnostic plot after-imputation.

Notes: Differences in the distribution of imputed data and the observed data for the urban variable. Imputation method *logreg.*



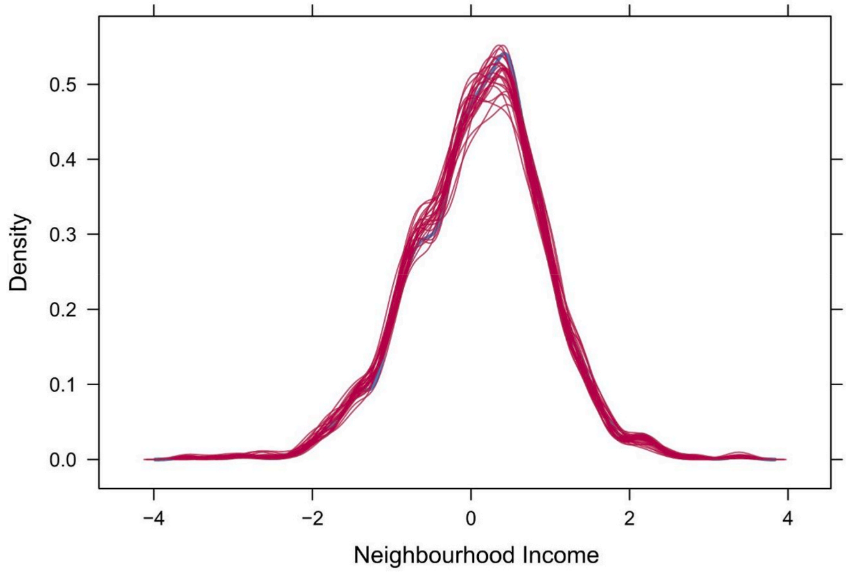


Fig. 5A. Diagnostic plot after-imputation.
Notes: Differences in the distribution of imputed data and the observed data for the neighbourhood income variable. Imputation method *pmm.*

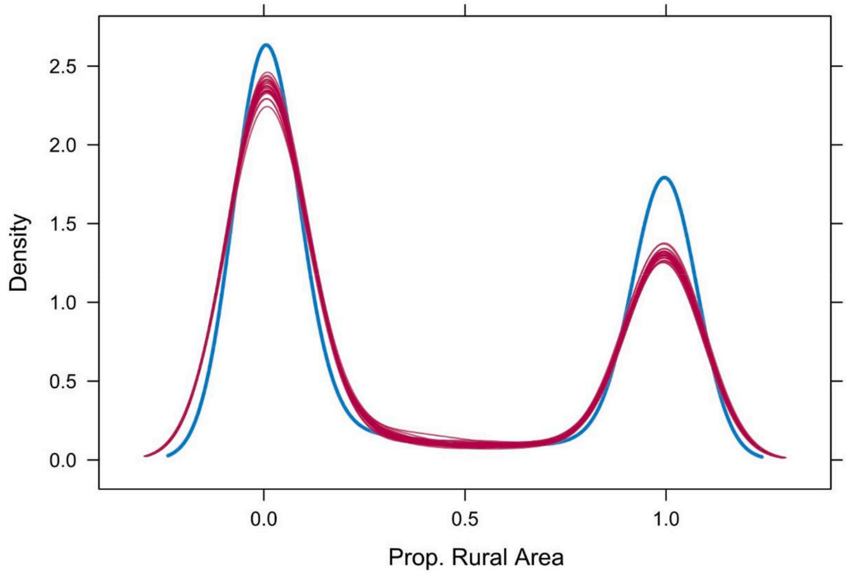


Fig. 5B. Diagnostic plot after-imputation.
Notes: Differences in the distribution of imputed data and the observed data for the proportion of rural area. Imputation method *pmm.*

Appendix B. Results from Sensitivity Analysis

Movers Variable

We created the variable of the movement across regions with the following method. While the Rural-Urban Commuting Area (RUCA) codes were available for Wave 1, they were absent for Wave 2. Thus, it would have been impossible to use solely the RUCA variable to capture the regional changes before high school.

To address the missing data in Wave 2, we employed a strategy using the continuous "Proportion Rural" variable available in both waves. We calibrated our definition of urban using the intersection of RUCA codes and population density in Wave 1. By analysing the empirical distribution of the proportion of rural population within the subsample classified as urban by RUCA, we identified a high concentration of residents in purely urban environments. Specifically, the 75th percentile of the distribution was located at a rural proportion of 0.01. This indicated that for the vast majority of respondents in administratively defined urban areas, the proportion of rural population was 1% or less. Consequently, to ensure strict comparability between Wave 1 and Wave 2, we established a conservative threshold of ≤ 0.01 (1%) rural population density to define the Urban context.

This threshold was applied identically to both waves to generate a binary environmental context variable.

- Urban: Proportion Rural ≤ 0.01
- Non-Urban/Rural: Proportion Rural ≥ 0.01

The final Mobility variable (movers) was constructed to identify respondents who experienced a shift in their living environment. A respondent was coded as a "Mover" (1) if their environmental classification changed between waves (i.e., transitioning from Urban to Non-Urban, or vice versa). Respondents who moved residence but remained within the same environmental classification (e.g., Urban to Urban) were coded as "Non-Movers" (0) regarding contextual exposure.

Table 1b
Replication with Neighbourhood Income (N = 4915)

Educational Attainment	(1)	(2)	(3)	(4)
PGI-EA	0.6167***	0.376***	0.338***	0.344**
	(0.044)	(0.035)	(0.054)	(0.054)
Family SES		0.826***	0.847***	0.843***
		(0.053)	(0.066)	(0.06)
Urban (ref: No) Yes		0.179**	0.176**	0.211*
		(0.065)	(0.06)	(0.079)
PGI-EA × Family SES			0.068	0.074
			(0.039)	(0.040)
PGI-EA × Urban				−0.061
				(0.062)
Family SES × Urban				0.076
				(0.048)
PGI-EA × Family SES × Urban				−0.005
				(0.047)
Control variables		X	X	X
Control interaction			X	X
Constant	14.145***	13.900***	13.876***	13.852***
	(0.087)	(0.072)	(0.073)	(0.075)
R-squared	0.062	0.258	0.265	0.272

Notes: Controlled for the PCs, sex, and age in all models. Further interactions controlled in model 4. Robust standard errors (SE) in parentheses; SE clustered by school classes; ***P < 0.001, **P < 0.01, *P < 0.05. All continuous independent variables are z-standardized prior to the analyses. Multiple imputation used to fill missing data (m = 25).

Table 2b
Replication with Urban 4 Categories (N = 4915)

Educational Attainment	(1)	(2)	(3)	(4)
PGI-EA	0.617***	0.385***	0.344***	0.221
	(0.044)	(0.035)	(0.055)	(0.132)
Family SES		0.908***	0.926***	0.980***
		(0.052)	(0.068)	(0.133)
RUCA (ref: Urban) Large, Rural City		−0.017	0.003	−0.407
		(0.213)	(0.06)	(0.079)
Small, Rural Town		0.171	0.162	−0.033
		(0.123)	(0.122)	(0.282)
Isolated Rural Town		0.362	0.350	0.172
		(0.211)	(0.213)	(0.108)
PGI-EA × Family SES			0.071	0.088
			(0.039)	(0.045)
PGI-EA × Large, Rural City				0.164
				(0.103)
× Small, Rural Town				−0.016
				(0.142)
× Isolated, Rural Town				0.172
				(0.108)
Family SES × Large, Rural City				−0.039
				(0.136)
× Small, Rural Town				−0.081
				(0.146)
× Isolated, Rural City				−0.398***
				(0.094)

(*continued on next page*)

Table 2b (*continued*)

Educational Attainment	(1)	(2)	(3)	(4)
PGI-EA × Family SES × Large, Rural City				0.015
				(0.152)
Small, Rural Town				−0.046
				(0.119)
Isolated, Rural Town				−0.114
				(0.117)
Control variables		X	X	X
Control interaction			X	X
Constant	14.145***	13.378***	13.357***	13.394***
	(0.087)	(0.136)	(0.132)	(0.153)
R-squared	0.062	0.257	0.263	0.280

Notes: Controlled for the PCs, sex, and age in all models. Further interactions controlled in model 4. Robust standard errors (SE) in parentheses; SE clustered by school classes; ***P < 0.001, **P < 0.01, *P < 0.05. All continuous independent variables are z-standardized prior to the analyses. Multiple imputation used to fill missing data (m = 25).

Table 3b

Regression Analysis Without MICE (N = 3573)

Educational Attainment	(1)	(2)	(3)	(4)
PGI-EA	0.626***	0.379***	0.362***	0.366**
	(0.046)	(0.037)	(0.057)	(0.094)
Family SES		0.916***	0.923***	0.778***
		(0.048)	(0.062)	(0.086)
Urban (ref: No) Yes		−0.254*	−0.252*	−0.107
		(0.120)	(0.122)	(0.163)
PGI-EA × Family SES			0.029	−0.019
			(0.042)	(0.076)
PGI-EA × Urban				−0.007
				(0.089)
Family SES × Urban				0.174
				(0.095)
PGI-EA × Family SES × Urban				0.057
				(0.092)
Control variables		X	X	X
Control interaction			X	X
Constant	14.164***	14.138***	14.116***	13.991***
	(0.085)	(0.113)	(0.117)	(0.139)
R-squared	0.054	0.239	0.237	0.264

Notes: Controlled for the PCs, sex, and age in all models. Further interactions controlled in model 4. Robust standard errors (SE) in parentheses; SE clustered by school classes; ***P < 0.001, **P < 0.01, *P < 0.05. All continuous independent variables are z-standardized prior to the analyses.

References

Abdellaoui, A., Dolan, C.V., Verweij, K.J.H., Nivard, M.G., 2022. Gene–environment correlations across geographic regions affect genome-wide association studies. Nat. Genet. 54 (9), 1345–1354. https://doi.org/10.1038/s41588-022-01158-0.

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K.E., Nivard, M.G., Veul, L., Holtz, Y., Zietsch, B.P., Frayling, T.M., Wray, N.R., Yang, J., Verweij, K.J.H., Visscher, P.M., 2019. Genetic correlates of social stratification in Great Britain. Nat. Hum. Behav. 3 (12), 1332–1342. https://doi.org/10.1038/s41562-019-0757-5.

Abdellaoui, A., Verweij, K.J.H., 2021. Dissecting polygenic signals from genome-wide association studies on human behaviour. Nat. Hum. Behav. 5 (6), 686–694. https://doi.org/10.1038/s41562-021-01110-y.

Adelman, C., 2006. The Toolbox Revisited. Paths to Degree Completion from High School Through College.

Alemu, R., Young, A.S., Benjamin, D.J., Turley, P., Okbay, A., 2025. Dissecting the predictive accuracy of polygenic indexes for behavioral phenotypes across genetic ancestries. Genetics. https://doi.org/10.1101/2025.09.11.675704.

Apple, M.W., Ball, S.J., Gandin, L.A., Di, A c, 2010. The Routledge International Handbook of the Sociology of Education. Routledge.

Arold, B., Hufe, P., Stoeckli, M., 2022. Genetic endowments, educational outcomes and the mediating influence of school investments. SSRN Electron. J. https://doi.org/10.2139/ssrn.4167444.

Autor, D., Figlio, D., Karbownik, K., Roth, J., Wasserman, M., 2016. School quality and the gender gap in educational achievement. Am. Econ. Rev. 106 (5), 289–295. https://doi.org/10.1257/aer.p20161074.

Ayorech, Z., Plomin, R., Von Stumm, S., 2019. Using DNA to predict educational trajectories in early adulthood. Dev. Psychol. 55 (5), 1088–1095. https://doi.org/10.1037/dev0000682.

Baier, T., Eilertsen, E.M., Ystrøm, E., Zambrana, I.M., Lyngstad, T.H., 2022a. An anatomy of the intergenerational correlation of educational attainment – learning from the educational attainments of Norwegian twins and their children. Res. Soc. Stratif. Mobil. 79, 100691. https://doi.org/10.1016/j.rssm.2022.100691.

Baier, T., Lang, V., 2019. The social stratification of environmental and genetic influences on education: new evidence using a register-based twin sample. Sociological Science 6, 143–171. https://doi.org/10.15195/v6.a6.

Baier, T., Lang, V., Grätz, M., Barclay, K.J., Conley, D.C., Dawes, C.T., Laidley, T., Lyngstad, T.H., 2022b. Genetic influences on educational achievement in cross-national perspective. Eur. Socio Rev. 38 (6), 959–974. https://doi.org/10.1093/esr/jcac014.

Banerjee, A.V., Cole, S., Duflo, E., Linden, L., 2007. Remedying education: evidence from two randomized experiments in India. Q. J. Econ. 122 (3), 1235–1264. https://doi.org/10.1162/qjec.122.3.1235.

Barone, C., Schizzerotto, A., Abbiati, G., Argentin, G., 2016. Information barriers, social inequality, and plans for higher education: evidence from a field experiment. Eur. Socio Rev. https://doi.org/10.1093/esr/jcw050 jcw050.

Behrman, J.R., Birdsall, N., 1983. The Quality of Schooling: Quantity Alone is Misleading.

Behrman, J.R., Mitchell, O.S., Soo, C., Bravo, D., 2010. Financial literacy, schooling, and wealth accumulation. SSRN Electron. J. https://doi.org/10.2139/ssrn.1688985.

Belsky, D., Domingue, B., Wedow, R., Arseneault, L., Boardman, J., Caspi, A., Conley, D., Fletcher, J.M., Freese, J., Herd, P., Moffitt, T., Poulton, R., Sicinski, K., Wertz, J., Harris, K., 2018. Genetic analysis of social-class mobility in five longitudinal studies. Proc. Natl. Acad. Sci. U. S. A 115, E7275–E7284. https://doi.org/10.1073/pnas.1801238115.

Belsky, D.W., Moffitt, T.E., Corcoran, D.L., Domingue, B., Harrington, H., Hogan, S., Houts, R., Ramrakha, S., Sugden, K., Williams, B.S., Poulton, R., Caspi, A., 2016. The genetics of success: how single-nucleotide polymorphisms associated with educational attainment relate to life-course development. Psychol. Sci. 27 (7), 957–972. https://doi.org/10.1177/0956797616643070.

Bernard, J., Steinführer, A., Klärner, A., Keim-Klärner, S., 2023. Regional opportunity structures: a research agenda to link spatial and social inequalities in rural areas. Prog. Hum. Geogr. 47 (1), 103–123. https://doi.org/10.1177/03091325221139980.

Bernardi, F., 2014. Compensatory advantage as a mechanism of educational inequality: a regression discontinuity based on month of birth. Sociol. Educ. 87 (2), 74–88. https://doi.org/10.1177/0038040714524258.

Bernardi, F., Triventi, M., 2020. Compensatory advantage in educational transitions: trivial or substantial? A simulated scenario analysis. Acta Sociol. 63 (1), 40–62. https://doi.org/10.1177/0001699318780950.

Biroli, P., Galama, T., Von Hinke, S., Van Kippersluis, H., Rietveld, C.A., Thom, K., 2025. The economics and econometrics of gene–environment interplay. Rev. Econ. Stud. https://doi.org/10.1093/restud/rdaf034 rdaf034.

Blackwell, D.L., McLaughlin, D.K., 1998. Do rural youth attain their educational goals? https://doi.org/10.22004/AG.ECON.289788.

Blanc, J., Mawass, W., Berg, J.J., 2025. Quantifying the susceptibility of polygenic scores to ancestry stratification. Genetics. https://doi.org/10.64898/2025.12.04.692430.

Boudon, R., 1974. Education, Opportunity, and Social Inequality: Changing Prospects in Western Society. Wiley.

Bourdieu, P., 1986. The forms of capital. In: Richardson, J. (Ed.), Handbook of Theory and Research for the Sociology of Education. Greenwood, New York, pp. 241–258.

Bourdieu, P., Passeron, J.-C., Bourdieu, P., Bourdieu, P., 1977. Reproduction in education, society and culture. In: Reprinted), 2. Sage Publ.

Branigan, A.R., McCallum, K.J., Freese, J., 2013. Variation in the heritability of educational attainment: an international meta-analysis. Soc. Forces 92 (1), 109–140. https://doi.org/10.1093/sf/sot076.

Breinholt, A., Conley, D., 2020. Child-Driven parenting: differential early childhood investment by offspring genotype. Soc. Forces 102, 310–329. https://doi.org/10.3386/w28217.

Brown, D.L., Schafft, K.A., 2018. Rural People and Communities in the 21st Century: Resilience and Transformation, 2 Edition. Polity Press.

Buchmann, C., DiPrete, T.A., McDaniel, A., 2008. Gender inequalities in education. Annu. Rev. Sociol. 34 (1), 319–337. https://doi.org/10.1146/annurev.soc.34.040507.134719.

Burt, C.H., 2024. Polygenic indices (aka polygenic scores) in social science: a guide for interpretation and evaluation. Sociol. Methodol. 54 (2), 300–350. https://doi.org/10.1177/00811750241236482.

Byun, S., Meece, J.L., Irvin, M.J., 2012. Rural-Nonrural disparities in Postsecondary educational Attainment revisited. Am. Educ. Res. J. 49 (3), 412–437. https://doi.org/10.3102/0002831211416344.

Cheadle, J.E., Amato, P.R., 2011. A quantitative assessment of Lareau's qualitative conclusions about class, race, and parenting. J. Fam. Issues 32 (5), 679–706. https://doi.org/10.1177/0192513x10386305.

Cheesman, R., Borgen, N.T., Lyngstad, T., Eilertsen, E., Ayorech, Z., Torvik, F., Andreassen, O., Zachrisson, H., Ystrøm, E., 2022. How genes, schools, and residential areas shape achievement: a population-wide gene-environment interaction study. https://www.semanticscholar.org/paper/273e275371a6c27ed0534caae63ade35544b4f95.

Cheesman, R., Borgen, N.T., Sandsør, A.M.J., Hufe, P., 2025. The Genetic Lottery Goes to School: Better Schools Compensate for the Effects of Students' Genetic Differences.

Cheesman, R., Hunjan, A., Coleman, J.R.I., Ahmadzadeh, Y., Plomin, R., McAdams, T.A., Eley, T.C., Breen, G., 2020. Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK biobank. Psychol. Sci. 31 (5), 582–591. https://doi.org/10.1177/0956797620904450.

Chen, P., 2014. Guidelines for analyzing add health data. https://doi.org/10.17615/C6BW8W.

Chetty, R., Friedman, J.N., 2010. Does Local Tax Financing of Public Schools Perpetuate Inequality?.

Chohan, B.I., Khan, R.M., 2010. Impact of Parental Support on the Academic Performance and Self Concept of the Student.

Coleman, J.S., 1988. Social capital in the creation of human capital. Am. J. Sociol.

Colodro-Conde, L., Rijsdijk, F., Tornero-Gómez, M.J., Sánchez-Romera, J.F., Ordoñana, J.R., 2015. Equality in educational Policy and the heritability of educational attainment. PLoS One 10 (11), e0143796. https://doi.org/10.1371/journal.pone.0143796.

De Zeeuw, E.L., De Geus, E.J.C., Boomsma, D.I., 2015. Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. Trends in Neuroscience and Education 4 (3), 69–76. https://doi.org/10.1016/j.tine.2015.06.001.

Dierker, P., Diewald, M., 2025. Compensation or accentuation? How parents from different social backgrounds decide to support their children. Eur. Socio Rev. 41 (1), 111–126. https://doi.org/10.1093/esr/jcae010.

Domingue, B., Trejo, S., Armstrong-Carter, E., Tucker-Drob, E., 2020. Interactions between polygenic scores and environments: methodological and conceptual challenges. Sociological Science 7, 365–386. https://doi.org/10.15195/v7.a19.

Domingue, B.W., Belsky, D.W., Conley, D., Harris, K.M., Boardman, J.D., 2015. Polygenic influence on educational attainment: new evidence from the National longitudinal Study of adolescent to adult health. AERA Open 1 (3). https://doi.org/10.1177/2332858415599972.

Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. Annu. Rev. Psychol. 53 (1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153.

Eifler, E.F., Riemann, R., 2022. The aetiology of educational attainment: a nuclear twin family study into the genetic and environmental influences on school leaving certificates. Br. J. Educ. Psychol. 92 (3), 881–897. https://doi.org/10.1111/bjep.12478.

Engzell, P., Tropf, F.C., 2019. Heritability of education rises with intergenerational mobility. Proc. Natl. Acad. Sci. 116 (51), 25386–25388. https://doi.org/10.1073/pnas.1912998116.

Erola, J., Lehti, H., Baier, T., Karhula, A., 2020. Socioeconomic background and gene-environment interplay in social stratification across the early life course. Eur. Socio Rev. https://doi.org/10.31235/osf.io/gahfc null, null.

Farrigan, T., Genetin, B., Sanders, A., Pender, J., Thomas, K.L., Winkler, R., Richelle, L.), Cromartie, J., United States. Department of Agriculture, Economic Research Service, 2024. Rural America at a Glance 2024 Edition. United States. Department of Agriculture, Economic Research Service. https://doi.org/10.32747/2024.8722498.ers.

Fertig, M., Wright, R.E., 2005. School quality, educational attainment and aggregation bias. Econ. Lett. 88 (1), 109–114. https://doi.org/10.1016/j.econlet.2004.12.028.

Fraemke, D., Willems, Y.E., Okbay, A., Lindenberger, U., Zinn, S., Wagner, G., Richter, D., Harden, K.P., Tucker-Drob, E.M., Hertwig, R., Koellinger, P., Raffington, L., 2025. Polygenic associations with educational attainment in East versus West Germany: differences emerge after reunification. Psychol. Sci. https://doi.org/10.1177/09567976251350965.

Ghirardi, G., Bernardi, F., 2025. Compensating or boosting genetic propensities? Gene-family socioeconomic status interactions by educational outcome selectivity. Soc. Sci. Res. 129, 103174. https://doi.org/10.1016/j.ssresearch.2025.103174.

Ghirardi, G., Gil-Hernández, C.J., Bernardi, F., Van Bergen, E., Demange, P., 2024. Interaction of family SES with children's genetic propensity for cognitive and noncognitive skills: no evidence of the Scarr-Rowe hypothesis for educational outcomes. Res. Soc. Stratif. Mobil. 92, 100960. https://doi.org/10.1016/j.rssm.2024.100960.

Giangrande, E.J., Beam, C.R., Carroll, S., Matthews, L.J., Davis, D.W., Finkel, D., Turkheimer, E., 2019. Multivariate analysis of the Scarr-Rowe interaction across middle childhood and early adolescence. Intelligence 77, 101400. https://doi.org/10.1016/j.intell.2019.101400.

Godwin, A.N., 2024. Add health as a resource for the science of the exposome: Rural-Urban commuting area (RUCA) codes. https://doi.org/10.17615/9YGQ-G064.

Halpern-Manners, A., Marahrens, H., Neiderhiser, J.M., Natsuaki, M.N., Shaw, D.S., Reiss, D., Leve, L.D., 2020. The intergenerational transmission of early educational advantages: new results based on an adoption design. Res. Soc. Stratif. Mobil. 67, 100486. https://doi.org/10.1016/j.rssm.2020.100486.

Harden, K.P., 2021. The Genetic Lottery: Why DNA Matters for Social Equality. Princeton University press.

Harel, O., 2009. The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. J. Appl. Stat. 36 (10), 1109–1118. https://doi.org/10.1080/02664760802553000.

Harerimana, N.V., Liu, Y., Ruks, M., 2025. Neighborhood context, genetic influences, and life satisfaction: evidence from the German twin family panel. PLoS One 20 (8), e0316416. https://doi.org/10.1371/journal.pone.0316416.

Hartung, A., Hillmert, S., 2019. Assessing the spatial scale of context effects: the example of neighbourhoods' educational composition and its relevance for individual aspirations. Soc. Sci. Res. 83, 102308. https://doi.org/10.1016/j.ssresearch.2019.05.001.

Haughbrook, R., Hart, S.A., Schatschneider, C., Taylor, J., 2016. Genetic and environmental influences on early literacy skills across school grade contexts. Dev. Sci. 20 (5), e12434. https://doi.org/10.1111/desc.12434.

Heath, A.C., Berg, K., Eaves, L.J., Solaas, M.H., Sundet, J., Nance, W.E., Corey, L.A., Magnus, P., 1985. No decline in assortative mating for educational level. Behav. Genet. 15 (4), 349–369. https://doi.org/10.1007/bf01070894.

Herd, P., Freese, J., Sicinski, K., Domingue, B., Harris, K.M., Wei, C., Hauser, R., 2019. Genes, gender inequality, and educational attainment. Am. Sociol. Rev. 84, 1069–1098. https://doi.org/10.1177/0003122419886550.

Hillman, N.W., Colston, J., Bach-Hanson, J., Peek, A., 2021. Mapping rural colleges and their communities. Am. Educ. Res. J. 53 (4), 987–1021. https://doi.org/10.3102/0002831216653204.

Hoff, E., 2006. How social contexts support and shape language development. Dev. Rev. 26 (1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002.

Howley, C.B., Gunn, E., 2003. Research About Mathematics Achievement in the Rural Circumstance.

Hummer, R.A., Hernandez, E.M., 2013. The effect of educational attainment on adult mortality in the United States. Popul. Bull. 68 (1), 1–16.

Jackson, M., 2013. Determined to succeed?: performance versus choice. In: Educational Attainment. Stanford University Press. https://doi.org/10.2307/j.ctvqsdrjr, 1a ed.

Kahneman, D., Deaton, A., 2010. High income improves evaluation of life but not emotional well-being. Proc. Natl. Acad. Sci. 107 (38), 16489–16493. https://doi.org/10.1073/pnas.1011492107.

Keller, M.C., 2014. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (Simple) solution. Biol. Psychiatry 75 (1), 18–24. https://doi.org/10.1016/j.biopsych.2013.09.006.

Knigge, A., Maas, I., Stienstra, K., Zeeuw, E. D. de, Boomsma, D., 2022. Delayed tracking and inequality of opportunity: Gene-environment interactions in educational attainment. NPJ Science of Learning 7. https://doi.org/10.1038/s41539-022-00122-1 null.

Kong, A., Thorleifsson, G., Frigge, M.L., Vilhjalmsson, B.J., Young, A.I., Thorgeirsson, T.E., Benonisdottir, S., Oddsson, A., Halldorsson, B.V., Masson, G., Gudbjartsson, D.F., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., Stefansson, K., 2018. The nature of nurture: effects of parental genotypes. Science 359 (6374), 424–428. https://doi.org/10.1126/science.aan6877.

Koricich, A., Chen, X., Hughes, R.P., 2018. Understanding the effects of rurality and socioeconomic status on college attendance and institutional choice in the United States. Rev. High. Educ. 41 (2), 281–305. https://doi.org/10.1353/rhe.2018.0004.

Krapohl, E., Rimfeld, K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.-B., Asbury, K., Harlaar, N., Kovas, Y., Dale, P.S., Plomin, R., 2014. The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. Proc. Natl. Acad. Sci. 111 (42), 15273–15278. https://doi.org/10.1073/pnas.1408777111.

Lahtinen, H., Korhonen, K., Martikainen, P., Morris, T., 2023. Polygenic prediction of education and its role in the intergenerational transmission of education: cohort changes among Finnish men and women born in 1925–1989. Demography 60 (5), 1523–1547. https://doi.org/10.1215/00703370-10963788.

Lareau, A., 2011. *Unequal Childhoods: Class, Race, and Family Life* (Second Edition, with an Update a Decade Later). University of California Press.

Lawrence, E., Hummer, R.A., Harris, K.M., 2017. The cardiovascular health of young adults: disparities along the urban-rural continuum. Ann. Am. Acad. Polit. Soc. Sci. 672 (1), 257–281. https://doi.org/10.1177/0002716217711426.

Lee, J.J., , COGENT (Cognitive Genomics Consortium), Social Science Genetic Association Consortium, Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M.A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P.N., 2018. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat. Genet. 50 (8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3.

Lin, M.-J., 2020. The social and genetic inheritance of educational attainment: genes, parental education, and educational expansion. Soc. Sci. Res. 86, 102387. https://doi.org/10.1016/j.ssresearch.2019.102387.

Liu, H., 2018. Social and genetic pathways in multigenerational transmission of educational attainment. Am. Sociol. Rev. 83 (2), 278–304. https://doi.org/10.1177/0003122418759651.

Marré, A., 2017. Rural Education at A Glance, 2017.

McCracken, J.D., Barcinas, J.D.T., 1991. Differences Between Rural and Urban Schools, Student Characteristics, and Student Aspirations in Ohio.

Mills, M., Tropf, F.C., 2020. Sociology, genetics, and the coming of Age of sociogenomics. Rev. Sociol. 46, 553–581. https://doi.org/10.1146/annurev-soc-121919-054756.

Morris, T.T., Davies, N.M., Davey Smith, G., 2020. Can education be personalised using pupils' genetic data? eLife 9. https://doi.org/10.7554/elife.49962.

Naylor, R.A., Smith, J., 2004. Determinants of educational success in higher education. Edward Elgar Publishing eBooks. https://doi.org/10.4337/9781845421694.00016.

Nivard, M., Belsky, D.W., Harden, K., Baier, T., Andreassen, O., Ystrøm, E., Bergen, E. van, Lyngstad, T., 2024. More than nature and nurture, indirect genetic effects on children's academic achievement are consequences of dynastic social processes. Nat. Hum. Behav. https://doi.org/10.1038/s41562-023-01796-2 null, null.

Nivard, M.G., Belsky, D.W., Harden, K.P., Baier, T., Ystrøm, E., Lyngstad, T.H., 2022. Neither Nature nor Nurture: Using Extended Pedigree Data to Elucidate the Origins of Indirect Genetic Effects on Offspring Educational Outcomes.

OECD, 2023. PISA 2022 Results (Volume I): the State of Learning and Equity in Education. OECD. https://doi.org/10.1787/53f23881-en.

Okbay, A., Beauchamp, J.P., Fontana, M.A., Lee, J.J., Pers, T.H., Rietveld, C.A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S.F.W., Oskarsson, S., Pickrell, J.K., Thom, K., Timshel, P., De Vlaming, R., Abdellaoui, A., Ahluwalia, T.S., Bacelis, J., Baumbach, C., et al., 2016. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533 (7604), 539–542. https://doi.org/10.1038/nature17671.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D.A., Ahlskog, R., Magnusson, P.K.E., Oskarsson, S., Hayward, C., Campbell, A., et al., 2022. Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nat. Genet. 54 (4), 437–449. https://doi.org/10.1038/s41588-022-01016-z.

Paasch, K.M., Swaim, P.L., 1995. More rural students are graduating from high school, but a serious dropout problem remains. https://doi.org/10.22004/AG.ECON.311074.

Pender, J., Hertz, T., Cromartie, J., Farrigan, T., 2019. Rural America at A Glance, 2019 Edition.

Pettersson, O., 2025. Raising the floor? Genetic influences on educational attainment through the lens of the evolving Swedish welfare State. Behav. Genet. 55 (3), 199–214. https://doi.org/10.1007/s10519-025-10219-z.

Pfeffer, F.T., 2008. Persistent inequality in educational attainment and its institutional context. Eur. Socio Rev. 24 (5), 543–565. https://doi.org/10.1093/esr/jcn026.

Plomin, R., 2019. Blueprint: How DNA Makes us who we are: with a New Afterword for the Paperback Edition (First MIT Press Paperback Edition). The MIT Press.

Plomin, R., Von Stumm, S., 2018. The new genetics of intelligence. Nat. Rev. Genet. 19 (3), 148–159. https://doi.org/10.1038/nrg.2017.104.

Polderman, T.J.C., Benyamin, B., De Leeuw, C.A., Sullivan, P.F., Van Bochoven, A., Visscher, P.M., Posthuma, D., 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat. Genet. 47 (7), 702–709. https://doi.org/10.1038/ng.3285.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38 (8), 904–909. https://doi.org/10.1038/ng1847.

Procopio, F., Liao, W., Rimfeld, K., Malanchini, M., Von Stumm, S., Allegrini, A.G., Plomin, R., 2025. Multi-polygenic score prediction of mathematics, reading, and language abilities independent of general cognitive ability. Mol. Psychiatr. 30 (2), 414–422. https://doi.org/10.1038/s41380-024-02671-w.

Ramos, R., Duque, J., Nieto, S., 2016. Decomposing the Rural-Urban Differential in Student Achievement in Colombia Using PISA Microdata, vols. 34–2, pp. 379–412. https://doi.org/10.25115/eae.v34i2.3044.

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J.R.I., Selzam, S., Dale, P.S., Esko, T., Metspalu, A., Plomin, R., 2018. Genetic influence on social outcomes during and after the Soviet era in Estonia. Nat. Hum. Behav. 2 (4), 269–275. https://doi.org/10.1038/s41562-018-0332-5.

Rodríguez-Pose, A., 2018. The revenge of the places that don't matter (and what to do about it). Camb. J. Reg. Econ. Soc. 11 (1), 189–209. https://doi.org/10.1093/cjres/rsx024.

Roscigno, V.J., Crowle, M.L., 2001. Rurality, institutional disadvantage, and Achievement/Attainment. Rural Sociol. 66 (2), 268–292. https://doi.org/10.1111/j.1549-0831.2001.tb00067.x.

Rowe, D.C., Jacobson, K.C., Van Den Oord, E.J.C.G., 1999. Genetic and environmental influences on vocabulary IQ: parental education level as moderator. Child Dev. 70 (5), 1151–1162. https://doi.org/10.1111/1467-8624.00084.

Rubin, D.B., 1987. Multiple Imputation for Nonresponse in Surveys. Wiley. https://doi.org/10.1002/9780470316696, 1^{a} ed.

Ruiz, J.D.C., Quackenboss, J.J., Tulve, N.S., 2016. Contributions of a child's built, natural, and social environments to their general cognitive ability: a systematic scoping review. PLoS One 11 (2), e0147741. https://doi.org/10.1371/journal.pone.0147741.

Ruks, M., 2022. Investigating the Mechanisms of G × SES Interactions for education. Res. Soc. Stratif. Mobil. https://doi.org/10.1016/j.rssm.2022.100730 null, null.

Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., Coleman, J.R.I., Grasby, K.L., Hammerschlag, A.R., Kaminski, J.A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C.A., Trampush, J.W., et al., 2018. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat. Genet. 50 (7), 912–919. https://doi.org/10.1038/s41588-018-0152-6.

Scarr-Salapatek, S., 1971. Race, Social Class, and IQ: population differences in heritability of IQ scores were found for racial and social class groups. Science 174 (4016), 1285–1295. https://doi.org/10.1126/science.174.4016.1285.

Sewell, W.H., Haller, A., Portes, A., 1969. The Educational and Early Occupational Attainment Process.

Showalter, D., Hartman, S., Johnson, J., Klein, B., 2019. Why rural matters 2018-2019: the time is now. A report of the rural school and community trust. https://eric.ed.gov/?id=ED604580.

Silva, L., Tropf, F.C., Wolfram, T., 2025. Mind the gap: the interplay between genes and neighbourhood context on educational achievement. Eur. Socio Rev. 41 (5), 755–773. https://doi.org/10.1093/esr/jcaf001.

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., Tomizawa, R., Watanabe, M., Sakai, N., Rebato, E., Busjahn, A., Tyler, J., Hopper, J., Ordoñana, J., Sánchez-Romera, J., Colodro-Conde, L., Calais-Ferreira, L., Oliveira, V.C., Ferreira, P., et al., 2020. Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. Sci. Rep. 10. https://doi.org/10.1038/s41598-020-69526-6 null.

Sirin, S.R., 2005. Socioeconomic status and academic achievement: a meta-analytic review of research. Rev. Educ. Res. 75 (3), 417–453. https://doi.org/10.3102/00346543075003417.

Stocké, V., Blossfeld, H.P., Hoenig, K., Sixt, M., 2019. Social inequality and educational decisions in the life course. In: Blossfeld, H.P., Roßbach, H.G. (Eds.), Education as a Lifelong Process. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-23162-0_6. Edition ZfE, vol. 3.

Spörlein, C., Zoch, G., Schlueter, E., 2025. Did social change strengthen genetic associations? Gendered educational attainment before and after German reunification. Res. Soc. Stratif. Mobil. 99, 101091. https://doi.org/10.1016/j.rssm.2025.101091.

Stienstra, K., Karlson, K., 2023. The nature-nurture of academic achievement at the intersection between gender, family background, and school context. Soc. Sci. Res. 111, 102870. https://doi.org/10.1016/j.ssresearch.2023.102870.

Stienstra, K., Knigge, A., Maas, I., 2024. Gene-environment interaction analysis of school quality and educational inequality. NPJ Science of Learning 9. https://doi.org/10.1038/s41539-024-00225-x null.

Tayyaba, S., 2012. Rural-urban gaps in academic achievement, schooling conditions, student, and teachers' characteristics in Pakistan. Int. J. Educ. Manag. 26 (1), 6–26. https://doi.org/10.1108/09513541211194356.

Tieken, M.C., 2014. Why Rural Schools Matter. University of North Carolina Press.

Trejo, S., Belsky, D., Boardman, J., Freese, J., Harris, K., Herd, P., Sicinski, K., Domingue, B., 2018. Schools as moderators of genetic associations with life course attainments: evidence from the WLS and add health. Sociological science 5, 513–540. https://doi.org/10.15195/v5.a22.

Tucker-Drob, E.M., Bates, T.C., 2016. Large cross-national differences in gene × socioeconomic status interaction on intelligence. Psychol. Sci. 27 (2), 138–149. https://doi.org/10.1177/0956797615612727.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., Gottesman, I.I., 2003. Socioeconomic status modifies heritability of IQ in young children. Psychol. Sci. 14 (6), 623–628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x.

Uchikoshi, F., Conley, D., 2021. Gene-environment interactions and school tracking during secondary education: evidence from the U.S. Res. Soc. Stratif. Mobil. 76. https://doi.org/10.1016/J.RSSM.2021.100628 null.

U.S. Bureau of the Census, 1992. 1990 Census of population and housing: population and housing unit counts (Report No. CPH-2). https://www.census.gov/library/publications/1992/dec/cph-2.html.

U.S. Department of Agriculture, Economic Research Service, 2025a. Rural-Urban Commuting Area Codes. U.S. Department of Agriculture. https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.

U.S. Department of Agriculture, Economic Research Service, 2025b. Rural-Urban Commuting Area Codes: Users' Guide. U.S. Department of Agriculture. https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/users-guide.

Van Buuren, S., 2018. Flexible Imputation of Missing Data, Second Edition. Chapman and Hall/CRC. https://doi.org/10.1201/9780429492259, 2^{a} ed.

Van De Werfhorst, H.G., Mijs, J.J.B., 2010. Achievement inequality and the institutional structure of educational systems: a comparative perspective. Annu. Rev. Sociol. 36 (1), 407–428. https://doi.org/10.1146/annurev.soc.012809.102538.

Van Hootegem, A., Rogne, A.F., Cros, C.T., Røgeberg, O., Lyngstad, T.H., 2025. Social origins and socioeconomic outcomes: a combined twin and adoption study. Eur. Socio Rev. 41 (5), 793–807. https://doi.org/10.1093/esr/jcaf029.

Wang, N., 2024. The urban–rural heterogeneous effect of family SES on achievement: the mediating role of culture. Behav. Sci. 14 (2), 84. https://doi.org/10.3390/bs14020084.

Weng, A.X., 2025. Education and mental health in young adulthood: new evidence from genetic markers. Health Econ. 34 (10), 1869–1881. https://doi.org/10.1002/hec.70012.

Weßling, K., Bechler, N., 2019. Where do regional influences matter? The impact of socio-spatial indicators on transitions from secondary school to university. Review of Regional Research 39 (2), 163–188. https://doi.org/10.1007/s10037-019-00131-5.

Williams, J.H., 2005. Cross-National Variations in Rural Mathematics Achievement: a Descriptive Overview.

Woodley Of Menie, M.A., Pallesen, J., Sarraf, M.A., 2018. Evidence for the Scarr–Rowe effect on genetic expressivity in a large U.S. sample. Twin Res. Hum. Genet. 21 (6), 495–501. https://doi.org/10.1017/thg.2018.63.

Young, A.I., 2022. Discovering missing heritability in whole-genome sequencing data. Nat. Genet. 54 (3), 224–226. https://doi.org/10.1038/s41588-022-01012-3.

Young, D.J., 1998. Rural and urban differences in student achievement in science and mathematics: a multilevel analysis. Sch. Effect. Sch. Improv. 9 (4), 386–418. https://doi.org/10.1080/0924345980090403.

Zahl-Thanem, A., Rye, J.F., 2024. Spatial inequality in higher education: a growing urban–rural educational gap? Eur. Socio Rev. 40 (6), 1067–1081. https://doi.org/10.1093/esr/jcae015.

Zhou, Q., Gidziela, A., Allegrini, A.G., Cheesman, R., Wertz, J., Maxwell, J., Plomin, R., Rimfeld, K., Malanchini, M., 2024. Gene-environment correlation: the role of family environment in academic development. Mol. Psychiatr. https://doi.org/10.1038/s41380-024-02716-0.