A Harmonic-Mean Decomposition of Sibling Correlations: Evidence from Turkey

Oğuz Öztunali*
Istanbul Bilgi University

Orhan Torul†
Boğaziçi University

April 2026

Abstract

We develop a decomposition framework that treats the sibling correlation—rather than its variance components—as the primitive object. The grand sibling correlation is exactly a weighted harmonic mean of within- and between-subset correlations; the identity is recursive and supports a counterfactual subset-removal exercise separating within- and between-channel contributions. Applying the framework to Turkish Demographic and Health Survey data for 1993–2018, we estimate an educational sibling correlation of 0.52 and document three findings. First, partitioning by parental education generates the largest between-subset amplification among the observable dimensions we examine, a level-shift analogue of Simpson's paradox in which every subset correlation lies below the grand correlation. Second, a sister–brother gap of 0.13 persists within every observable partition of family background and geography. Third, subset-removal effects operate primarily through reweighting in the harmonic aggregation rather than through changes in correlation levels—a mechanism that additive variance decompositions cannot uncover.

Keywords: sibling correlations; educational mobility; harmonic mean decomposition; Simpson's paradox; intergenerational transmission; Turkey; Demographic and Health Survey.

JEL codes: J62, I24, O15, J13, D31.

*Address: Istanbul Bilgi University, Department of Economics, 34060 Eyüp, Istanbul, Turkey. E-Mail: oguz.oztunali@bilgi.edu.tr

†Corresponding author. Address: Boğaziçi University, Department of Economics, 34342 Bebek, Istanbul, Turkey. E-Mail: orhan.torul@bogazici.edu.tr

1 Introduction

Sibling correlations summarize the importance of shared family background for individual outcomes by measuring the share of total variation attributable to factors common to siblings raised in the same family. Because siblings share parental characteristics as well as neighborhood, school, and genetic influences, the sibling correlation is a more comprehensive measure of family background than the intergenerational elasticity, which captures only the direct parent-child transmission of a specific resource. Cross-country estimates vary substantially—generally higher in the United States and Germany than in the Nordic countries—but evidence for developing countries remains limited, in part because parent-child linked data are scarce in low- and middle-income settings. Demographic and Health Survey (DHS)–style co-resident samples provide one of the few vehicles for measurement, and Turkey is a useful laboratory: a rapidly modernising middle-income economy with persistent educational inequalities along gender, geography, and parental lines.

A recent strand of the literature has begun to address the measurement gap by decomposing sibling correlations across observable dimensions of family background. Karlson and In (2025), Karlson (2025), and Hällsten (2026) apply the law of total variance to the between-family variance component of the intra-class correlation, partitioning its numerator into additive within- and between-group pieces while treating the total variance as a common normalizing constant. These approaches provide valuable accounting decompositions of the variance component underlying the sibling correlation, but they do not characterize the aggregation structure of the correlation coefficient itself. Because the sibling correlation is a ratio of additively decomposable sample moments, its aggregation follows a weighted harmonic mean rather than a weighted average—not by assumption, but as an algebraic consequence of the ratio structure, provided the between-subset cross-term between the family and idiosyncratic components is negligible (a condition we verify empirically for every partition).

In this paper, we develop a decomposition framework that takes the sibling correlation itself as the primitive object of analysis. We show that the grand sibling correlation can be expressed exactly as a weighted harmonic mean of (i) the weighted harmonic mean of within-subset sibling correlations and (ii) the between-subset sibling correlation, with variance-based weights that capture the relative importance of each channel. The resulting framework has three features that distinguish it from existing approaches. First, it characterizes the coefficient-level relationship linking the grand, within-, and between-subset sibling correlations without imposing assumptions on how the variance components behave across contexts. Second, because within- and between-subset sibling correlations share the same mathematical structure, the decomposition can be applied recursively, allowing researchers to progressively refine their partitioning while preserving the exact relationship to the population-wide measure at each stage. Third, it supports a counterfactual subset-removal exercise that cleanly separates two channels through which a given subset contributes to the grand correlation: the reshaping of the harmonic mean of the remaining subsets, and the alteration of the between-subset structure.

We apply this framework to Turkish Demographic and Health Survey data and obtain several findings of substantive interest. Our estimates place the educational sibling correlation in Turkey at approximately 0.52 for the full sample, above the corresponding estimates for Denmark, Norway, and Sweden but below those for the United States and Germany. Two features of the Turkish case stand out. First, partitioning the sample by parental education generates the largest between-subset amplification among all observable dimensions we examine: every subset-specific sibling correlation lies below the grand correlation, a compositional aggregation effect that we interpret as a level-shift analogue of Simpson's paradox. Second, the sister-only subsample exhibits an educational sibling correlation of 0.67, which is 0.13 points higher than the brother-only estimate—a gap that is substantially larger than those typically reported for developed countries and whose direction contrasts with the United States and Germany, though it matches the pattern reported for Denmark by Bredtmann and Smith (2018). Applying our decomposition recursively, we show that this sister-brother gap, as well as a related gender-based firstborn effect we document, persists within every observable partition of family characteristics we consider, pointing to unobserved gender-specific mechanisms that the standard observables do not capture. Finally, our counterfactual exercise reveals a reweighting mechanism through which subset removal affects the grand sibling correlation primarily by changing the geometry of the harmonic aggregation rather than the underlying correlation levels—a channel that additive variance decompositions cannot, by construction, detect.

These findings bear directly on inequality-of-opportunity measurement. Sibling correlations are widely interpreted as summary measures of circumstance inequality in the Roemer sense, so a framework that localizes *which* observable circumstance contributes to the aggregate through *which* channel—within-subset persistence versus between-subset stratification—has immediate application for policy targeting. In the Turkish case, our decomposition identifies parental education as the dominant stratifying dimension, indicates that the sister-brother gap is produced by unobserved family- and community-level factors rather than by any of the standard observables, and shows that removing particular subgroups reshapes the aggregate correlation mainly by altering the weight structure of the harmonic aggregator rather than by moving within- or between-subset correlations themselves. Each of these is a statement about inequality of opportunity that additive variance decompositions cannot, in general, make.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on sibling correlations and recent decomposition approaches. Section 3 describes the DHS data and develops our harmonic mean decomposition framework. Section 4 presents the estimation results, the single-layer decomposition and counterfactual exercise, and the dual-layer decomposition of the gender-based gaps. Section 5 concludes.

2 Literature Review

Solon (1999) formally demonstrates that the intergenerational elasticity is a component of the sibling correlation, so the latter nests the former while also capturing shared influences such as neighborhood effects, school quality, and genetic endowments (Björklund and Jäntti, 2009). Björklund and Jäntti (2020) show that these non-intergenerational factors make up the bulk of sibling correlations in Sweden and the United States: the intergenerational elasticity accounts for only 21% and 35% of the sibling correlation in earnings, respectively. The contemporary estimation framework dates back to Solon et al. (1991), who model the unobserved determinants of a socioeconomic outcome as the sum of an orthogonal family-specific and idiosyncratic component, and recover their variances via an ANOVA-based approach. Mazumder (2008) later proposes a REML-based alternative that relaxes homoskedasticity and accommodates different distributional forms.

In this study, we estimate sibling correlations with respect to years of education in the Turkish context for which previous findings are limited. In contrast, educational sibling correlations have been extensively studied in developed countries. Educational sibling correlations tend to be higher in the United States and Germany compared to the Nordic countries, consistent with the broader pattern observed for earnings. Mazumder (2008) estimates an educational sibling correlation of 0.602 in the pooled sample, 0.622 in the brothers-only sample, and 0.602 in the sisters-only sample in the United States. In the German context, Schnitzlein (2014) reports a sibling correlation of 0.656 for brothers and 0.551 for sisters. Bredtmann and Smith (2018) report an educational sibling correlation of 0.38 for mixed sex siblings, 0.39 for brothers, and 0.42 for sisters with respect to years of education in Denmark. For Norway, Björklund and Salvanes (2011) estimate a sibling correlation of 0.40 for all siblings, 0.42 for brothers, and 0.39 for sisters with respect to years of education. Björklund and Jäntti (2012) report sibling correlations of 0.40 for mixed sexes, 0.46 for brothers, and 0.40 for sisters based on years of schooling in the Swedish context.

To our knowledge, the only studies reporting educational sibling correlations for Turkey are the recent contributions by Ahsan et al. (2023) and Muñoz and Jaque (2026). Using Demographic and Health Survey (DHS) data for 53 developing countries, Ahsan et al. (2023) estimate an overall educational sibling correlation of 0.547 for Turkey—slightly above our full-sample estimate of 0.52, a difference attributable to their use of the Bingley and Cappellari (2019) estimator, their inclusion of all co-resident siblings rather than two-oldest-sibling dyads, and differences in sample-construction criteria—and document a modest inverse U-shaped pattern over time, with the estimate rising from 0.550 for the 1970s cohort to 0.561 for the 1980s cohort before declining to 0.433 for the 1990s cohort. Furthermore, following the decomposition methodology proposed by Bingley and Cappellari (2019) to quantify the intergenerational component of the sibling correlation, they find that the intergenerational channel accounts for a substantial 77% of the educational sibling correlation. This figure is not directly comparable to the earnings-based 21%/35% estimates of Björklund and Jäntti (2020), since it refers to a different outcome and uses the Bingley and Cappellari (2019) decomposition methodology rather than the traditional regression-based approach. Using the 1985, 1990, and 2000 waves of the IPUMS data and imputing years of schooling from categorical education information, Muñoz and Jaque (2026) report educational sibling correlations for Turkey in the range of 0.54–0.55.

Our analysis also connects to work investigating the correlates of sibling correlations. Björklund et al. (2010) estimate how sibling correlations respond to the inclusion of indicators of socio-economic status (such as parental income, education, and social class), family structure (such as mother's age at first birth, parents' cohabitation and marital status), and social problems (such as parents' presence, mental health, alcoholism) in Sweden. Forsberg et al. (2025) investigate how sibling correlations vary with the parental socioeconomic ventile (based on income or education) and find non-linear and non-monotonic associations. Anger and Schnitzlein (2017) report sibling correlations conditional on family income and maternal education in Germany. Using a novel genetic dataset in the UK, Fletcher et al. (2023) find that molecular genetics has only a modest impact on sibling correlations, confirming that shared socioeconomic environments rather than genetic similarity are the primary driver.

Our paper is also closely related to the newly emerging literature on the decomposition of sibling correlations. The recent exchange between Hällsten (2026), Karlson and In (2025), and Karlson (2025) addresses heterogeneity in sibling correlations by decomposing the intra-class correlation (ICC) understood as a variance ratio. Karlson and In (2025) apply the law of total variance to partition the numerator additively into between-group mean deviations and within-group family variances, Hällsten (2026) uses this to show that split-sample ICCs condition out mean differences and can reverse substantive conclusions, and Karlson (2025) argues that group-specific ICCs cannot separately identify the transmission parameter from the dispersion of latent family factors — a fundamental identification problem. All three operate within the variance-component framework, decomposing the numerator of the ICC while treating total variance as a normalizing constant.

Our approach differs from the existing decomposition approaches structurally: we work with the sibling correlation itself as the primitive object rather than its variance components. In doing so, we take the sibling correlation as a descriptive summary of the joint distribution of sibling outcomes rather than as a reduced-form expression of an underlying transmission parameter. The identification critique of Karlson (2025) therefore does not apply to our framework in the same way: we do not claim to separately identify the transmission mechanism from latent dispersion, but instead characterize how the sibling correlation—whatever its structural origin—aggregates across observable partitions. Because the sibling correlation is a ratio of two additively decomposable sample moments, its aggregation follows a weighted harmonic mean — an algebraic consequence of the ratio structure rather than an imposed assumption. The resulting decomposition yields the coefficient-level functional form through which the population-wide sibling correlation depends on within-group and between-group correlations. Since these objects share the same mathematical structure at every level of aggregation, the decomposition can be applied recursively — partitioning groups into finer subgroups and decomposing again — thereby allowing researchers to progressively zoom into the data while preserving the precise relationship to the population-wide measure at each stage. This nested harmonic structure captures the non-linear interaction between within-group similarity and between-group stratification, an aspect that the additive Karlson–In framework does not characterize at the coefficient level. Our counterfactual $\ddot{\Delta}_m$ term exploits precisely this non-additivity to measure each group's contribution operating through changes in the between-group structure. In this sense, our framework complements the variance-component approach by revealing the aggregation structure that governs the relationship between group-level and population-level sibling correlations.

3 Data and Methodology

3.1 Data

In our analysis, we rely on the Turkish microdata collected by Hacettepe University as a part of the Demography and Health Surveys (DHS) Program overseen by USAID. We utilize the entire data available in the 1993, 1998, 2003, 2008, 2013, and 2018 waves of the survey. Following Ahsan et al. (2023), who uses the DHS dataset to estimate educational sibling correlations for 53 countries, and Ahsan et al. (2025), we focus on individuals aged 16 to 28 in order to minimize the impact of co-residency bias since our observations belong to siblings residing in the same survey unit. Although we follow Mazumder (2008) and Solon et al. (1991) rather than the Bingley and Cappellari (2019) estimator used by Ahsan et al. (2023), we adopt their sample-construction conventions to maintain comparability. To this end, we construct family units consisting of two parents and their two oldest children. In families with more than two children, we leave out the remaining younger siblings in order (i) for families with larger sizes not to have a disproportionate effect on our estimates and (ii) to keep the findings of the decomposition methodology we develop to study how sibling correlations interact with observable shared family variables in a tractable manner.

Our sample consists of siblings born between 1970 and 1999. Since birth cohort will be one of the correlates of educational sibling correlations we will be exploring, we define three pure and two mixed-birth cohort profiles based on siblings' decade of birth. The 1970s, 1980s, and 1990s pure cohort profiles correspond to the sibling dyads where both siblings are born in 1970-1979, 1980-1989, and 1990-1999 periods, respectively. The two mixed-cohort sibling duos are then captured by the 1970s-1980s and 1980s-1990s mixed-cohort profiles where the older and younger siblings are born in two consecutive decades.[1] We keep all observations for which we have complete information regarding the correlates of educational sibling correlations we explore, and based on the resulting inclusion criteria, we end up with 7,892 unique sibling duos and a total of 15,784 observations.

One of the main goals of our study is to explore the variation in educational sibling correlations based on family-specific variables, namely father's and mother's educational attainment[2] as well as the age at which their eldest child was born, the number of siblings in the family, as well as the type of children's current place of residence.[3]

In Table 1, we report summary statistics for the entire sample as well as the subsets of data defined on the

[1]While technically we could have a 1970s-1990s mixed-cohort profile in our analysis, due to the fact that we are limiting our scope to sibling duos with a maximum potential age difference of 12 years, as a result of our exclusion criteria we do not have any observations corresponding to this cohort profile.

[2]In DHS the level of educational attainment is recorded in four categories: (i) 'no education' refers to the lack of any formal education, (ii) 'primary' stands for primary school education, (iii) 'secondary' encompasses middle and high school education, and (iv) 'higher' corresponds to tertiary education.

[3]As we do not have information regarding whether the child's place of residence was rural or not, we opted for current place of residence for which we have information in DHS.

subcategories of these variables. When we explore the educational attainment trends across the subcategories of our correlates, we first observe that there is a sharp increase in mean years of schooling across birth cohorts, while the standard deviation first rises from the 1970s to the 1980s before declining in the 1990s. Two patterns stand out: (i) a positive association between children's mean years of schooling and their parents' educational attainment, and (ii) a negative relationship between the standard deviation of children's years of schooling and parental educational attainment. For example, while the mean and standard deviation of the years of education of children with uneducated fathers are 6.10 and 3.70, we observe a mean and a standard deviation of 12.33 and 2.52 for children with highly educated fathers. Furthermore, there exists a non-linear, inverse U-shaped association between the age at which the father had his eldest child and children's educational attainment. While the average years of education of children was 7.72 in families where the eldest child was born when the father was younger than age 20, this average rises to 9.22 for fathers in their 20s, declines to 8.43 for fathers in their 30s, and falls further to 7.06 for fathers aged 40 or above when the eldest sibling was born. The pattern for mothers is similar in direction but flatter at younger ages, declining more sharply from the 30s onward.

Table 1: Descriptive Statistics

Variable: Years of Education Completed		Both Siblings		Older Sibling		Younger Sibling		Father		Mother		# of Families
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	
Whole Sample		8.82	3.76	9.01	4.01	8.63	3.49	5.59	3.85	3.08	3.37	7892
Birth Cohorts	1970s	7.61	3.63	7.74	3.80	7.48	3.44	4.58	3.64	2.30	3.01	2122
	1970–1980s	8.35	3.84	8.61	4.08	8.09	3.57	5.28	4.01	2.82	3.35	1008
	1980s	8.69	3.86	8.85	4.11	8.53	3.60	5.88	3.90	3.13	3.48	2251
	1980–1990s	9.63	3.26	9.82	3.71	9.44	2.73	6.13	3.72	3.62	3.42	887
	1990s	10.43	3.32	10.72	3.56	10.14	3.03	6.41	3.73	3.87	3.42	1624
Father's Education	No education	6.10	3.70	6.10	3.83	6.09	3.56	0.00	0.00	0.51	1.42	1305
	Primary	8.47	3.51	8.58	3.72	8.37	3.28	4.78	0.79	2.61	2.65	4301
	Secondary	10.66	2.95	11.05	3.12	10.26	2.70	9.05	1.87	4.60	3.34	1770
	Higher	12.33	2.52	13.06	2.60	11.60	2.21	14.67	1.36	8.25	4.34	516
Mother's Education	No education	7.31	3.71	7.35	3.87	7.27	3.56	3.75	3.11	0.00	0.00	3479
	Primary	9.60	3.36	9.84	3.60	9.36	3.07	6.24	3.19	4.51	1.17	3617
	Secondary	11.74	2.63	12.36	2.77	11.12	2.32	10.03	3.86	9.15	2.03	659
	Higher	12.61	2.49	13.41	2.56	11.81	2.15	13.83	2.85	14.15	1.93	137
Sibling Gender Composition	Brothers	8.70	3.45	8.84	3.69	8.56	3.19	5.24	3.62	2.85	3.31	2900
	Sisters	8.90	4.16	9.00	4.52	8.81	3.76	6.19	4.16	3.57	3.56	1336
	Male/Female	8.79	3.82	9.48	3.57	8.09	3.94	5.54	3.82	3.01	3.25	2092
	Female/Male	9.02	3.88	8.72	4.58	9.32	2.98	5.80	3.96	3.17	3.43	1564
Sibling Count	Two siblings	9.27	3.53	9.55	3.80	9.00	3.23	6.11	3.89	3.67	3.52	5142
	Three siblings	8.32	3.92	8.33	4.11	8.30	3.73	4.90	3.65	2.27	2.90	1912
	Four or more siblings	7.18	4.15	7.26	4.33	7.10	3.97	3.98	3.33	1.24	2.23	838
Degree of Urbanization	Urban	9.51	3.63	9.79	3.87	9.24	3.35	6.29	4.03	3.62	3.58	5336
	Rural	7.38	3.63	7.40	3.82	7.35	3.43	4.15	2.97	1.95	2.55	2556
First Fatherhood Age	Below 20s	7.72	3.65	7.79	3.84	7.64	3.45	5.42	2.69	2.29	2.56	273
	20s	9.22	3.61	9.45	3.88	8.99	3.31	6.26	3.63	3.61	3.39	5008
	30s	8.43	3.90	8.60	4.10	8.26	3.69	4.82	4.07	2.42	3.30	2112
	40s or above	7.06	3.90	7.07	4.14	7.05	3.65	2.26	3.16	0.91	2.16	499
First Motherhood Age	Below 20s	8.91	3.59	9.18	3.83	8.65	3.31	5.99	3.58	3.28	2.93	1679
	20s	9.09	3.72	9.30	3.99	8.88	3.43	6.00	3.84	3.40	3.51	5013
	30s	7.61	3.91	7.64	4.06	7.57	3.76	3.46	3.53	1.51	2.89	1115
	40s or above	6.94	3.95	6.98	4.11	6.89	3.82	2.05	2.62	0.40	1.29	85
Sibling Age Difference	0 to 1 year	8.59	3.69	8.71	3.88	8.48	3.50	5.70	3.81	3.11	3.24	1687
	2 years	8.71	3.75	8.87	3.93	8.55	3.56	5.36	3.68	2.77	3.19	2353
	3 years	9.01	3.77	9.24	3.97	8.78	3.55	5.64	3.93	3.18	3.46	1451
	4 years	8.86	3.85	9.18	4.11	8.54	3.56	5.64	3.96	3.16	3.37	945
	5 years or more	9.05	3.78	9.27	4.24	8.82	3.24	5.76	4.02	3.38	3.67	1456

Notes: Birth-cohort categories are mutually exclusive and defined by both siblings' decades of birth: "1970s," "1980s," and "1990s" denote dyads in which both siblings were born in the same decade, while "1970s–1980s" and "1980s–1990s" denote dyads in which the older and younger siblings were born in two consecutive decades.

One of the most striking educational attainment patterns pertains to the gender composition of the sibling dyad. Ranking the four dyad types by mean years of schooling yields brother–brother dyads lowest at 8.70, followed by male-firstborn mixed-gender dyads at 8.79, sister–sister dyads at 8.90, and female-firstborn mixed-gender dyads highest at 9.02. Within brother-only dyads, the younger sibling is slightly less educated than the older one, and within both mixed-gender dyads the male sibling is more educated than the female sibling regardless of birth order. There is also a negative (positive) relationship between the mean (standard deviation) of children's years of education and the number of siblings they have. With respect to the age difference between siblings, both the older and younger sibling's mean years of schooling increase with the gap, but the standard deviation of the older sibling's years of education rises while that of the younger sibling remains relatively stable. Finally, as expected, there is a considerable advantage of siblings who are currently residing in urban residences while the standard deviation of years of schooling is nearly identical across urban and rural areas (3.63 in both), though the unconditional standard deviation (3.76) is noticeably higher, reflecting the compositional effect of the between-group mean difference.

Sample construction and limitations. Four features of the sample shape the interpretation of our estimates. First, because DHS records co-resident siblings only, our estimates are subject to co-residency bias: siblings who leave the household—disproportionately the more educated and geographically mobile—are missing from the sample, and the remaining co-resident pairs tend to be more similar than a random pair from the same family. The direction of the resulting bias is generally expected to be upward. Because early marriage and rural-to-urban migration affect young women's departure from the household more than young men's in Turkey, the co-residency selection may also differ by gender, which is relevant for interpreting the sister-brother gap. The age restriction to 16–28 is a life-cycle window in which co-residency is still the norm for most young adults in Turkey, which attenuates but does not eliminate the bias; Ahsan et al. (2025) provides a detailed discussion. Second, individuals in the 16–22 range may still be enrolled in school, so their completed years of education are right-censored. This compresses variation most heavily for the 1990s birth cohort and may contribute mechanically to the cohort decline we document in Section 4.1. Third, restricting attention to the two oldest children narrows the within-family age gap relative to a random-pair design, which may modestly bias the sibling correlation upward. Our data, however, provide a clean diagnostic: among families with exactly two children, the "two oldest" rule does not bind, because those two children are the only two. As reported in Table 2, this subsample (5,142 dyads, 65% of the sample) yields an educational sibling correlation of 0.52, statistically indistinguishable from the full-sample estimate of 0.52 and from the three-child subsample estimate of 0.51. The pair-selection restriction therefore does not appear to drive the main result; the sharper decline observed in families with four or more children is consistent with genuine differences in family environments rather than with the pair-selection rule itself. Fourth, as a consequence of the first three, cross-country comparisons with the Nordic register data or US and German administrative data should be read as indicative rather than as direct quantitative benchmarking.

3.2 Methodology

3.2.1 Estimation

Following Solon et al. (1991), we model the years of schooling of sibling j of family i, i.e. Y_{ij}, as follows:

$$Y_{ij} = \Omega + \phi X_{ij} + \epsilon_{ij} \tag{1}$$

where Ω stands for the constant term, X_{ij} corresponds to a vector of individual-specific control variables (a set of indicator variables controlling for the birth year of the respondent and the year in which the relevant wave's data was collected in our case), and ϕ refers to the vector of fixed coefficients (i.e., individual-invariant) associated with control variables. ϵ_{ij} stands for the composite error term which is modeled as shown in (2):

$$\epsilon_{ij} = \alpha_i + u_{ij} \tag{2}$$

where α_i is a family-specific random coefficient that is identical for all siblings in family i with $\alpha_i \sim (0, \sigma^2_\alpha)$ $\forall i \in 1, \ldots, N$, and u_{ij} is an idiosyncratic error term with $u_{ij} \sim (0, \sigma^2_u)$ $\forall i \in 1, \ldots, N$ and $\forall j \in 1, \ldots, J_i$ where

J_i[4] stands for the number of siblings in family i and N corresponds to the total number of families. Assuming the family-specific and idiosyncratic components are orthogonal, $\mathbb{E}(\alpha_i \cdot u_{ij}) = 0$, and that idiosyncratic errors are uncorrelated across siblings within the same family, $\mathbb{E}(u_{ij} \cdot u_{ik}) = 0$ for $j \neq k$, the variance of the composite error term can be expressed as the sum of the variance of the family-specific component σ_α^2 and that of the idiosyncratic component σ_u^2 as displayed by (3):

$$\sigma_\epsilon^2 = \sigma_\alpha^2 + \sigma_u^2 \tag{3}$$

In this setting, the educational sibling correlation ρ is the intra-class correlation of the composite error terms ϵ_{ij}—equivalently, the share of the variance of ϵ_{ij} that is accounted for by the family-specific component σ_α^2:

$$\rho = \frac{\sigma_\alpha^2}{\sigma_\epsilon^2} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} \tag{4}$$

where the residual $\hat{\epsilon}_{ij}$ is obtained from the fitted regression in (1):

$$\hat{\epsilon}_{ij} = Y_{ij} - \hat{\Omega} - \hat{\phi} X_{ij}. \tag{5}$$

In our estimations, to demean siblings years of education and obtain the educational sibling correlation ρ for the entire sample and conditional on observable family traits, we mainly use the Restricted Maximum Likelihood (REML) based approach of Mazumder (2008), which additionally assumes normality of α_i and u_{ij}, while estimating (1). For robustness purposes, we also repeat our estimations using Solon et al. (1991)'s ANOVA based estimation strategy and display our findings in the Appendix.

3.2.2 Decomposition

Beyond estimating educational sibling correlations for the full sample and for subsets defined by observable shared family traits, we aim to (i) quantify the factors driving differences across subset-specific educational sibling correlations, and (ii) express the sample-wide sibling correlation as a function of within-subset correlations to assess how different groups shape the aggregate.

The derivation proceeds in four steps. We first contrast our objective with the two most common heterogeneity strategies in the literature—random-slope REML specifications and iterative regression—and motivate working with the sibling correlation as the primitive object. We next decompose the cross-sectional variance of each unobservable into within- and between-subset components, which leads to the central identity of the paper: the grand sibling correlation equals the weighted harmonic mean of the within- and between-subset correlations, not their weighted average. We then exploit the recursive structure of this identity to decompose subset-level differences into a within- and between-subsubset channel. Finally, we derive a counterfactual subset-removal exercise that separates the two mechanisms through which a given subset shapes the grand correlation.

There are different approaches one can take to study how educational sibling correlations vary with observed shared family traits. One is the strategy adopted by Mazumder (2008) and Bredtmann and Smith (2018), which estimates an augmented REML specification with random slopes for observed family traits,

[4] While we use a flexible notation for the number of siblings in a family, we set $J_i = 2$ in our empirical analysis for the reasons discussed in the data description section. The decomposition derived in Appendix A holds for arbitrary J_i: the cross-sectional variance decomposition in (A.6) – (A.8) and the harmonic mean expression in (A.18) make no use of $J_i = 2$ and go through unchanged for any family-size distribution, balanced or unbalanced. The estimation methodologies of Solon et al. (1991) and Mazumder (2008) likewise accommodate variable family sizes. The $J_i = 2$ restriction is therefore a sample-construction choice rather than a methodological constraint.

thereby allowing heterogeneity in sibling correlations to be related partly to observables and partly to unobserved shared family factors. This strategy imposes a strong functional-form assumption on how observed and unobserved family traits interact, and the resulting heterogeneity is hard to interpret: as emphasized by Karlson (2025), subgroup differences within such latent-variable transmission models may reflect differences in the transmission parameter, in the dispersion of the latent family factor, or in idiosyncratic noise, and these cannot be separately identified. Our framework sidesteps this identification problem by taking the sibling correlation itself as the descriptive object of interest and characterizing its exact aggregation structure across observable partitions, rather than attempting to recover a latent transmission parameter separately from latent dispersion and idiosyncratic noise. When the observable family characteristics are categorical, as in our case, the random-slope interpretation is also less natural, since it amounts to assigning a family-specific return to a categorical indicator.

Another approach, used by Björklund et al. (2010), is to add observed shared family characteristics as regressors and examine how the estimated sibling correlation changes across specifications. A potential limitation of this strategy is that, in addition to relying on functional-form assumptions about the interaction between observed and unobserved family traits, it can become difficult to implement and interpret in multivariate settings. In particular, introducing multiple related shared family characteristics may generate multicollinearity, sparse cells, or near-saturation, especially when several of the relevant observables are categorical. Our decomposition approach circumvents this difficulty by allowing multivariate analysis through nested decompositions.

Finally, when it works, the iterative regression approach quantifies only the reduction in the sibling correlation associated with the introduction of observed family characteristics. By contrast, our decomposition approach can recover that type of diagnostic information without the limitations of the foregoing methods, while also revealing the structure of the sibling correlation itself. In particular, it decomposes the aggregate correlation into within- and between-subset channels and shows how different subgroup partitions shape the aggregate through each of them, without reducing the family effect itself. More broadly, our results suggest caution in interpreting regression-based reductions in the sibling correlation after conditioning on a socioeconomic variable as measuring that variable's distinct contribution. Because the sibling correlation is generated through a non-linear aggregation of within- and between-subset structures, the meaning of such reductions is not straightforward.

In this sense, our decomposition approach is not a substitute for regression-based strategies but a complement to them: while the latter ask how much of the family effect can be attributed to observables, ours asks how the family effect is structured across observable dimensions.

To this end, we first partition our dataset into K mutually exclusive groups and define the educational sibling correlation of the k$^{\text{th}}$ subset, namely ρ_k, in the following way:

$$\rho_k = \frac{\sigma^2_{\alpha,k}}{\sigma^2_{\epsilon,k}} = \frac{\sigma^2_{\alpha,k}}{\sigma^2_{\alpha,k} + \sigma^2_{u,k}} \tag{6}$$

where $\sigma^2_{\epsilon,k}$, $\sigma^2_{\alpha,k}$, and $\sigma^2_{u,k}$ refer to the within-subset variance of the composite error term, the family-specific error term, and the idiosyncratic error term, respectively. ρ_k stands for the share of family-specific unobservables in the total variance of the unobserved factors influencing the educational attainment of an individual in subset k. The definition carries the orthogonality assumption of the underlying model subset-wise: $\mathbb{E}(\alpha_{i,k} \cdot u_{i,j,k}) = 0$ for every subset k, which is inherited from the specification in (2) applied within each subset.

Next, in the Appendix, inspired by the approach used by Hertz (2008) to decompose the intergenerational regression and correlation coefficients cross-sectionally, we show that any sample-wide variance term can be additively decomposed into a within-subset and between-subset component in a cross-sectional manner in the following way:

$$\sigma_\alpha^2 = \underbrace{\sum_{k=1}^{K} \gamma_k \sigma_{\alpha,k}^2}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}_{\substack{\text{between} \\ \text{component}}} \tag{7}$$

$$\sigma_u^2 = \underbrace{\sum_{k=1}^{K} \gamma_k \sigma_{u,k}^2}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{u}_k - \overline{u})^2}_{\substack{\text{between} \\ \text{component}}} \tag{8}$$

$$\sigma_\epsilon^2 = \underbrace{\sum_{k=1}^{K} \gamma_k \sigma_{\epsilon,k}^2}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{\epsilon}_k - \overline{\epsilon})^2}_{\substack{\text{between} \\ \text{component}}} \tag{9}$$

where $\gamma_k \equiv \frac{n(k)-1}{[\sum_{k=1}^{K} n(k)-1]}$ and $\eta_k \equiv \frac{n(k)}{[\sum_{k=1}^{K} n(k)-1]}$ are the weights of the kth subset in the within and between components of the respective cross-sectional variance decomposition. In this decomposition, the *within components* are truly within in the sense that they are just the cross-sectional weighted sums of subset-specific analogues of the relevant variance term. We address the *between components* as between since they are only defined over subset means and the grand sample mean and just correspond to the cross-sectional sums of the distances between these two means. [5]

The feature that distinguishes the *between component* from the *within component* is that while the latter is just a convex combination of subset variances and is a *variance-based* object, the former is not. Instead, the *between component* is a *variance-like* object defined over subsets. It resembles a variance since it is a weighted sum of the deviations of subset means from the grand mean. However, unlike a variance where observations have equal weights, in this object, subsets have an influence proportional to their sizes. Moreover, the normalizing denominator of the between component, inherited from the sum-of-squares identity in (7)–(9), is not $K-1$ but $\sum_{k=1}^{K} n(k) - 1$, so the object is not a standard between-group variance in the ANOVA sense; we use this denominator throughout purely as a normalizing convention rather than as an inferential degrees-of-freedom statement.

After the cross-sectional decomposition of the variance terms and the introduction of the $within$ and $between$ components, we now define the following between-subset sibling correlation $\tilde{\rho}$:

$$\tilde{\rho} = \frac{\tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\epsilon^2} = \frac{\tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_u^2} \tag{10}$$

[5]Since the estimation methodologies used in our study can only identify the variance of the family-specific component and not its levels, we can retrieve σ_α^2 residually via (7) as "between component $= \sigma_\alpha^2 -$ within component". This can occur when the between component is retrieved residually with overlapping subsets (overlapping in the sense of intersection α distributions across subsets) or having close $\bar{\alpha}_k$ values across subsets as the $\gamma_k \sigma_{\alpha,k}^2$ terms overshoot and yield a negative between component - which is a phenomenon we have observed while using the variance components estimated via REML by Mazumder (2008).

where

$$\tilde{\sigma}^2_\alpha = \sum_{k=1}^{K} \eta_k(\overline{\alpha}_k - \overline{\alpha})^2 \tag{11}$$

$$\tilde{\sigma}^2_u = \sum_{k=1}^{K} \eta_k(\overline{u}_k - \overline{u})^2 \tag{12}$$

$$\tilde{\sigma}^2_\epsilon = \sum_{k=1}^{K} \eta_k(\overline{\epsilon}_k - \overline{\epsilon})^2 \tag{13}$$

where $\tilde{\sigma}^2_\epsilon$, $\tilde{\sigma}^2_\alpha$, and $\tilde{\sigma}^2_u$ are simply the between components obtained after the cross-sectional decomposition of the respective variance terms. $\tilde{\rho}$ can be interpreted as the share of family-specific unobservables in the total variance of the unobserved factors influencing individuals' educational attainment across groups. Two small technical remarks are in order. First, the additive identity $\tilde{\sigma}^2_\epsilon = \tilde{\sigma}^2_\alpha + \tilde{\sigma}^2_u$ requires the weighted between-subset covariance of the subset means $\overline{\alpha}_k$ and $\overline{u}_k$, $\sum_k \eta_k(\overline{\alpha}_k - \overline{\alpha})(\overline{u}_k - \overline{u})$, to vanish. The individual-level orthogonality $\mathbb{E}(\alpha \cdot u) = 0$ does not mechanically guarantee this in finite samples; in our data the cross-term is numerically negligible for every partition we consider, and the additive split holds to close approximation throughout. Second, while u has zero family-conditional mean by construction, its between-subset component $\sum_k \eta_k(\overline{u}_k - \overline{u})^2$ captures any residual covariation between subset assignment and the idiosyncratic component—for instance, cohort or survey-wave effects not absorbed by X—which is empirically small in our setting.

After some algebra, in the Appendix, we prove that the sample-wide sibling correlation ρ can be expressed as a function of the weighted harmonic mean of the within-group correlation coefficients $\hat{\rho}$ and the between-group correlation coefficient $\tilde{\rho}$ as follows:[6]

$$\rho = \frac{V + W}{\dfrac{V}{\hat{\rho}} + \dfrac{W}{\tilde{\rho}}} \tag{14}$$

where $v_k \equiv \gamma_k \sigma^2_{\alpha,k}$ and $w_k \equiv \eta_k(\overline{\alpha}_k - \overline{\alpha})^2$ are the variance-weighted contributions of the k$^{\text{th}}$ group, $V \equiv \sum_{k=1}^{K} v_k$, $W \equiv \sum_{k=1}^{K} w_k$, and

$$\hat{\rho} \equiv \frac{V}{\displaystyle\sum_{k=1}^{K} \frac{v_k}{\rho_k}} \tag{15}$$

is the variance-weighted harmonic mean of the subset-specific sibling correlations ρ_k, with weights v_k. As (14) shows, holding the weights V and W fixed, the sample-wide correlation ρ is an increasing, non-linear function of both the weighted harmonic mean of within-subset correlations $\hat{\rho}$ and the between-subset correlation $\tilde{\rho}$. In order to quantify the impact of the difference between $\hat{\rho}$ and $\tilde{\rho}$ on ρ, we define the relative deviation of ρ from $\hat{\rho}$:

$$\pi = \frac{\rho - \hat{\rho}}{\hat{\rho}} = \frac{W(\tilde{\rho} - \hat{\rho})}{V\tilde{\rho} + W\hat{\rho}} \tag{16}$$

and express the sample-wide correlation as

$$\rho = \hat{\rho}(1 + \pi) \tag{17}$$

which describes the sample-wide grand educational sibling correlation as the multiplication of two components: (i) the weighted harmonic mean of within-subset correlations $\hat{\rho}$, and (ii) the relative deviation of the

[6]See Appendix for the definitions of the within-group and between-group variance components and the detailed description of our decomposition approach.

grand correlation from $\hat{\rho}$ - i.e., $1+\pi$. Holding π constant, the sample-wide grand educational sibling correlation is proportional to $\hat{\rho}$. Therefore, we refer to $\hat{\rho}$ as the *within-effect* on the grand correlation. Furthermore, as (15) clearly demonstrates, π is positively related to $\tilde{\rho}-\hat{\rho}$ and quantifies how the discrepancy between $\hat{\rho}$ and $\tilde{\rho}$ pulls ρ toward $\tilde{\rho}$ - which we refer to as the *between-effect* on the grand educational sibling correlation. Equation (16) also implies a trade-off between $\hat{\rho}$ and π across partitioning schemes applied to the same sample. Because the grand sibling correlation ρ is fixed for a given sample, a partitioning scheme that yields a larger weighted harmonic mean of subset correlations must necessarily yield a smaller between-effect. That is, if two partitioning schemes A and B are applied to the same sample and $\hat{\rho}_A > \hat{\rho}_B$, then it must follow that $\pi_A < \pi_B$.

The π term in (16) is helpful in summarizing the relative contributions of the components of ρ. As the absolute value of π approaches zero, the role of cross-sectional differences, hence of the between correlation, vanishes, and ρ increasingly approaches the weighted harmonic mean of the within-subset correlations. This occurs because, as π approaches zero, the within-, between-, and grand correlations converge toward one another, so the distinction between within- and between-group structure becomes progressively weaker. Note that any nonzero π indicates that the aggregate correlation differs from the within harmonic mean — a generalized form of Simpson's paradox. The sign of π determines the direction: positive π indicates the between-group structure amplifies the aggregate beyond the within effect, while negative π indicates suppression.

Although the aggregation rule linking ρ to $\hat{\rho}$ and $\tilde{\rho}$ is formally non-linear, the empirical importance of the between-subset amplification term may vary across applications. In cases where π is small, the grand sibling correlation can lie numerically close to the weighted harmonic mean of the within-subset sibling correlations, even though the exact aggregation structure remains harmonic rather than additive.

A simple numerical illustration clarifies what the harmonic structure reveals that an additive variance-component decomposition cannot. Consider two equal-sized subsets with $\rho_1 = \rho_2 = 0.30$ and a between-subset correlation $\tilde{\rho} = 0.90$, with variance-based weights $V = W = 1$. The grand correlation is

$$\rho = \frac{V+W}{V/\hat{\rho} + W/\tilde{\rho}} = \frac{2}{1/0.30 + 1/0.90} = 0.45,$$

which lies above every subset-specific correlation (here $\rho_1 = \rho_2 = 0.30 < 0.45$)—a level-shift analogue of Simpson's paradox that arises whenever $\tilde{\rho}$ is sufficiently large relative to the ρ_k. An additive variance decomposition of σ_α^2 into within- and between-subset components would correctly report that, under this parameterization, half of σ_α^2 is attributable to the between-subset channel. That statement is accurate as an accounting identity but is silent on two questions of direct substantive interest: by how much does the between channel pull the aggregate correlation away from the within structure, and how would that pull respond to a change in the subset configuration? The harmonic form answers both directly. The between-subset amplification is summarized by $\pi = (0.45 - 0.30)/0.30 = 0.50$. If $\tilde{\rho}$ rises from 0.90 to 0.95 while weights are held fixed, ρ moves only to 0.456; if, instead, W rises by fifty percent with $\tilde{\rho}$ held fixed, ρ moves to 0.500. Equal-sized changes on the two sides of the formula produce very different effects on the aggregate. This asymmetry between levels and weights is a structural feature of harmonic aggregation, and it is what the coefficient-level framework is designed to expose. Additive decompositions attribute variance; the harmonic framework attributes *pull on the aggregate correlation*, and these are different questions.

Our decomposition approach differs from the existing decomposition approaches proposed by Karlson and In (2025) and Karlson (2025)[7]. Those approaches start from the additive decomposition of the between-

[7] Hällsten (2026) is a recent study that builds on the decomposition approach of Karlson and In (2025) in order to address the attenuation of sibling correlations conditional on high parental income reported by Forsberg et al. (2025), by constructing composite and between-group ICCs based on global-mean centering.

family variance component, σ^2_α, via the law of total variance, while treating the population variance in the denominator of the ICC as a common normalizing constant. As a result, they provide an accounting decomposition of the variance component underlying the sibling correlation, rather than an aggregation theorem for the correlation coefficient itself.[8]

Our approach instead takes seriously the fact that the sibling correlation is a ratio-valued object, so that its numerator and denominator are algebraically linked across grand, within-, and between-subset contexts. Exploiting the parallel functional forms of the within- and between-subset sibling correlations, we derive a weighted harmonic mean decomposition of the grand sibling correlation in terms of within- and between-subset correlations, without imposing assumptions on how the variance components behave across contexts. Rather than imposing an additive decomposition on a fundamentally non-additive object, our approach directly characterizes the coefficient-level relationship linking the grand sibling correlation to its within- and between-subset counterparts. This relationship holds regardless of whether the data are balanced or unbalanced. In our framework, the heterogeneous roles of the within and between components are summarized by two objects only: $\hat{\rho}$, the weighted harmonic mean of subset-specific sibling correlations, and π, the Simpson's paradox coefficient.

The weighted-harmonic mean expression of sibling correlations reveals a recursive structure. If K subsets are each partitioned into L finer subsets along a second dimension, then the k$^{\text{th}}$ subset's educational sibling correlation itself can be expressed as a weighted harmonic mean of (i) the weighted harmonic mean of its subsubset correlations $\hat{\rho}_k$, and (ii) the between-subsubset correlation $\tilde{\rho}_k$:

$$\rho_k = \frac{V_k + W_k}{\dfrac{V_k}{\hat{\rho}_k} + \dfrac{W_k}{\tilde{\rho}_k}} \tag{18}$$

where $v_{k,l} \equiv \gamma_{k,l}\sigma^2_{\alpha,k,l}$ and $w_{k,l} \equiv \eta_{k,l}(\overline{\alpha}_{k,l} - \overline{\alpha}_k)^2$ are the second-level variance-weighted contributions of the l$^{\text{th}}$ subsubset of the k$^{\text{th}}$ subset—exact parallels of v_k and w_k in (14), now centered on the subset mean $\overline{\alpha}_k$ rather than the grand mean $\overline{\alpha}$—with $V_k \equiv \sum_{l=1}^{L} v_{k,l}$ and $W_k \equiv \sum_{l=1}^{L} w_{k,l}$. Or, more compactly:

$$\rho_k = \hat{\rho}_k(1 + \pi_k) \tag{19}$$

Using (18), the absolute difference between the sibling correlations of subsets s and r can be written as

$$\rho_s - \rho_r = \underbrace{(1 + \pi_r)(\hat{\rho}_s - \hat{\rho}_r)}_{\substack{\text{within-subsubset} \\ \text{effect}}} + \underbrace{\hat{\rho}_s(\pi_s - \pi_r)}_{\substack{\text{between-subsubset} \\ \text{effect}}}. \tag{20}$$

The first term may be interpreted as the *within-subsubset effect*, since it captures the part of the difference attributable to differences in the weighted harmonic means of the two subsets' subsubset correlations, holding the between-effect fixed at the level of subset r. The second term may be interpreted as the *between-subsubset effect*, since it captures the additional part of the difference attributable to the difference in the between-effects of the two subsets. Intuitively, the *within-subsubset effect* reflects the impact of differences in the weighted harmonic mean structure, whereas the *between-subsubset effect* reflects the impact of differences in the between-subsubset structure across the L subsubsets of the two subsets. As in any Oaxaca–Blinder-type decomposition, the choice of reference group is not innocuous: exchanging s and r yields a symmetric decomposition with within and between terms rearranged accordingly. We report results for a single reference throughout for expositional simplicity; symmetric decompositions can be recovered as a Shapley average of the two order-

[8] A direct numerical comparison between our coefficient-level decomposition and the variance-component decomposition of Karlson and In (2025) is left to future work, since the two frameworks decompose different objects (the correlation coefficient versus the between-family variance component) and the mapping between their respective summary quantities is not one-to-one.

ings and do not change our substantive conclusions. The recursive nature of our approach therefore makes it possible to investigate why subsets differ from one another. Once a given subset appears distinct at one level of aggregation, the same decomposition can be applied within that subset to determine whether its difference is driven by internal heterogeneity, by finer between-subgroup stratification, or by the interaction of the two.

Finally, our harmonic mean decomposition allows us to quantify the impact of the removal of a subset on the grand sibling correlation. To this end, let ρ^c_{-m} be the counterfactual sibling correlation that arises when the between-subset correlation $\tilde{\rho}$ is held constant while the weighted harmonic mean of the remaining $K-1$ subsets changes and becomes $\hat{\rho}_{-m} = \frac{\sum_{-m} v_k}{\sum_{-m} \frac{v_k}{\rho_k}}$ as a result of this exercise:

$$\rho^c_{-m} = \frac{V_{-m} + W_{-m}}{\frac{V_{-m}}{\hat{\rho}_{-m}} + \frac{W_{-m}}{\tilde{\rho}}} \tag{21}$$

where $V_{-m} = \sum_{-m} v_k = V - v_m$ and $W_{-m} = \sum_{-m} w_k = W - w_m$. The counterfactual ρ^c_{-m} and the actual grand correlation resulting from the removal of the m^{th} subset can be expressed as:

$$\rho^c_{-m} = \hat{\rho}_{-m}(1 + \pi^c_{-m}) \tag{22}$$

$$\rho_{-m} = \hat{\rho}_{-m}(1 + \pi_{-m}) \tag{23}$$

Letting Δ_m denote the absolute change from ρ to ρ_{-m}, one can express it as the sum of the following two components:

$$\Delta_m = \rho_{-m} - \rho = \underbrace{\left[\hat{\rho}_{-m}(1 + \pi^c_{-m}) - \hat{\rho}(1 + \pi)\right]}_{\equiv \dot{\Delta}_m} + \underbrace{\hat{\rho}_{-m}(\pi_{-m} - \pi^c_{-m})}_{\equiv \ddot{\Delta}_m} \tag{24}$$

where $\dot{\Delta}_m$ stands for the impact of the change in the harmonic mean structure of the remaining subsets, while $\ddot{\Delta}_m$ corresponds to the effect of the change in the between-subset structure after the removal of the m^{th} subset. This counterfactual exercise based on our decomposition strategy reveals how much of the change in ρ is due to (i) the change in the harmonic mean structure of the remaining subsets while isolating the impact of the change in the between-subset structure: $\dot{\Delta}_m$ captures the change from $\hat{\rho}$ to $\hat{\rho}_{-m}$ while holding the between component constant at $\tilde{\rho}$, thereby yielding a counterfactual Simpson's paradox coefficient π^c_{-m} for the remaining subsets; and (ii) the change in the between-subset structure, captured by $\ddot{\Delta}_m$, while isolating it from the impact of the change in the harmonic mean structure of the remaining subsets. Intuitively, this exercise shows how the removed subset m contributes to ρ through two channels: $\dot{\Delta}_m$ captures how removing m reshapes the harmonic mean of the remaining subsets, while $\ddot{\Delta}_m$ captures how the removal alters the between-subset structure. In this sense, large values of $\ddot{\Delta}_m$ indicate that the relevant partitioning scheme is carving the data into markedly different groups, thereby revealing that the associated socioeconomic dimension is an important axis of inequality or dissimilarity in the sample. More broadly, through $\ddot{\Delta}_m$ and π, our decomposition approach may also prove useful for policy discussions conducted under financial constraints by helping identify which socioeconomic partitions generate the strongest between-group amplification of sibling correlations and, more generally, which dimensions of stratification are most closely associated with limited educational mobility.

Note that, unlike Karlson and In (2025), our counterfactual analysis is not based on hypothetical changes to variance components treated as independently manipulable objects. Instead, it is a diagnostic decomposition that quantifies the two channels through which a particular subset affects the grand sibling correlation. We follow this route because the non-linear dependence of the grand sibling correlation on its within- and between-subset components implies that the effect of changing any one component necessarily interacts with the others. Our exercise therefore isolates contribution channels defined by the coefficient-level aggregation structure itself, rather than imposing hypothetical variations on objects whose effects cannot, in general, be cleanly separated. Moreover, unlike additive variance decompositions, our framework does not support statements of the form that a given subset "accounts for $x\%$" of the sibling correlation. Such language presumes separable additive contributions. In our setting, subsets affect the aggregate correlation through non-linear interactions operating via the within- and between-subset channels, so their importance is better understood in terms of structural influence than as percentage shares.

4 Results

We first present the educational sibling correlations for the full sample and for subsets defined by observed shared family characteristics, based on the REML estimator of Mazumder (2008). We then decompose the grand educational sibling correlation along these dimensions to quantify how observed shared family characteristics shape the grand correlation through the within and between channels. Finally, we apply the decomposition recursively to account for the gender-specific patterns revealed by our estimates.

4.1 Educational Sibling Correlation Estimates

Table 2 reports the educational sibling correlation estimates together with the variance of the family-specific component, σ^2_α, and the idiosyncratic component, σ^2_u, for the full sample and for subsets defined by observed shared family characteristics. Table C.1 presents the corresponding ANOVA-based estimates of Solon et al. (1991). The REML- and ANOVA-based estimates are highly similar, suggesting that the main empirical patterns are not driven by the choice of estimation method.

As discussed in Section 3.1, our estimates are based on co-resident sibling dyads, whereas the comparison estimates from the Nordic countries, the United States, and Germany are typically obtained from register or administrative data that include siblings regardless of co-residency status. The cross-country comparison below should therefore be interpreted as indicative rather than as a direct quantitative benchmarking.

Our REML estimates yield an educational sibling correlation of 0.52 for the full sample, which includes brothers, sisters, and mixed-gender sibling dyads. The corresponding estimate is 0.54 for brothers, 0.67 for sisters, 0.51 for mixed-gender dyads in which the firstborn is male, and 0.44 for mixed-gender dyads in which the firstborn is female. Relative to the findings discussed in the literature review, these estimates place Turkey above Denmark, Norway, and Sweden, but below the United States and Germany. One particularly notable feature of the Turkish case, however, is the magnitude of the gender gap in educational sibling correlations. In Turkey, the sister-only subsample exhibits an educational sibling correlation 0.13 points higher than that of the brother-only subsample—a gap that is substantially larger than those typically reported for developed countries. Schnitzlein (2014) documents a gap of comparable magnitude for Germany, but with the opposite sign (0.656 for brothers versus 0.551 for sisters); Mazumder (2008) reports essentially no gender gap in the United States. The direction of the Turkish gap matches the Danish estimates of Bredtmann and Smith (2018) (0.42 for sisters versus 0.39 for brothers), but its magnitude is roughly four times larger.[9]

[9] This pattern is consistent with a broader literature documenting that educational investment in daughters in Turkey is more tightly tied to family circumstances than investment in sons, for whom schooling is more normatively expected regardless of family background. Rankin and Aytaç (2006), Dayioğlu (2005), Smits and Hoşgör (2006), and Kırdar et al. (2018) document substantial heterogeneity in female schooling outcomes by family socioeconomic status, parental education, and region, while male schooling outcomes are comparatively more uniform. In Öztunalı and Torul (2022), using the Turkish Statistical Institute's Intergenerational Transmission of Disadvantages Module for 2011, we likewise report higher intergenerational educational correlation coefficients for females than for males across cohorts born between 1951 and 1985, lending further support to this interpretation. Under such a pattern, families play a larger role in determining sisters' educational attainment than brothers', which would manifest as a higher family-specific variance component and therefore a higher sibling correlation among sisters. We return to this point in Section 4.3

Table 2: Educational sibling correlations and variance components by subgroup based on estimates via Mazumder (2008)

Dimension	Group	σ^2_α Estimate [95% CI]	σ^2_u Estimate [95% CI]	σ^2_ϵ Estimate [95% CI]	ρ Estimate [95% CI]
Whole sample		6.698 [6.446, 6.950]	6.068 [5.873, 6.263]	12.766 [12.502, 13.030]	0.525 [0.511, 0.539]
Birth cohorts					
	1970s	6.997 [6.556, 7.437]	5.870 [5.544, 6.197]	12.867 [12.433, 13.301]	0.544 [0.519, 0.568]
	1970–1980s	7.378 [6.712, 8.044]	7.242 [6.691, 7.793]	14.620 [13.918, 15.322]	0.505 [0.471, 0.539]
	1980s	8.330 [7.824, 8.836]	6.348 [6.012, 6.683]	14.677 [14.183, 15.171]	0.568 [0.545, 0.590]
	1980–1990s	5.376 [4.658, 6.094]	5.144 [4.611, 5.677]	10.520 [9.707, 11.333]	0.511 [0.466, 0.556]
	1990s	4.315 [3.791, 4.840]	5.628 [5.125, 6.132]	9.944 [9.289, 10.599]	0.434 [0.393, 0.475]
Father's education					
	No education	4.837 [4.239, 5.436]	8.119 [7.556, 8.681]	12.956 [12.316, 13.596]	0.373 [0.335, 0.411]
	Primary	5.042 [4.755, 5.328]	6.118 [5.863, 6.373]	11.160 [10.847, 11.473]	0.452 [0.432, 0.472]
	Secondary	2.872 [2.526, 3.218]	4.756 [4.419, 5.092]	7.628 [7.196, 8.060]	0.377 [0.341, 0.412]
	Higher	0.983 [0.563, 1.402]	3.368 [2.858, 3.877]	4.350 [3.603, 5.098]	0.226 [0.160, 0.292]
Mother's education					
	No education	5.539 [5.182, 5.897]	7.331 [7.020, 7.643]	12.871 [12.495, 13.247]	0.430 [0.409, 0.452]
	Primary	4.636 [4.349, 4.923]	5.362 [5.101, 5.622]	9.997 [9.673, 10.321]	0.464 [0.441, 0.486]
	Secondary	2.616 [2.117, 3.116]	2.915 [2.607, 3.224]	5.532 [4.953, 6.111]	0.473 [0.422, 0.524]
	Higher	0.397 [-0.784, 1.579]	3.037 [2.071, 4.003]	3.434 [1.662, 5.206]	0.116 [-0.067, 0.299]
First fatherhood age					
	Below 20s	5.688 [4.256, 7.119]	7.052 [5.973, 8.131]	12.740 [11.280, 14.199]	0.446 [0.366, 0.527]
	20s	6.147 [5.846, 6.448]	5.712 [5.485, 5.940]	11.859 [11.531, 12.187]	0.518 [0.501, 0.536]
	30s	6.983 [6.460, 7.506]	6.577 [6.183, 6.971]	13.559 [13.005, 14.114]	0.515 [0.488, 0.542]
	40s or above	6.876 [5.824, 7.928]	7.157 [6.414, 7.899]	14.033 [13.003, 15.063]	0.490 [0.439, 0.542]
First motherhood age					
	Below 20s	5.828 [5.366, 6.291]	5.770 [5.419, 6.121]	11.598 [11.099, 12.097]	0.503 [0.475, 0.530]
	20s	6.569 [6.262, 6.875]	5.920 [5.680, 6.160]	12.489 [12.153, 12.824]	0.526 [0.509, 0.543]
	30s	7.005 [6.347, 7.662]	6.887 [6.361, 7.414]	13.892 [13.177, 14.607]	0.504 [0.470, 0.538]
	40s or above	5.790 [0.994, 10.586]	7.932 [6.343, 9.522]	13.722 [9.265, 18.180]	0.422 [0.293, 0.551]
Sibling gender composition					
	Brothers	6.038 [5.680, 6.395]	5.079 [4.819, 5.340]	11.117 [10.741, 11.493]	0.543 [0.521, 0.565]
	Sisters	10.066 [9.299, 10.834]	4.936 [4.513, 5.358]	15.002 [14.190, 15.814]	0.671 [0.644, 0.698]
	Male/Female	6.592 [6.099, 7.084]	6.231 [5.879, 6.583]	12.823 [12.317, 13.329]	0.514 [0.488, 0.540]
	Female/Male	5.842 [5.293, 6.392]	7.548 [7.113, 7.983]	13.390 [12.759, 14.021]	0.436 [0.408, 0.464]
Sibling count					
	Two	5.605 [5.342, 5.869]	5.139 [4.925, 5.354]	10.745 [10.448, 11.041]	0.522 [0.504, 0.539]
	Three	7.292 [6.802, 7.781]	7.054 [6.641, 7.467]	14.346 [13.836, 14.856]	0.508 [0.482, 0.534]
	Four or more	7.039 [6.101, 7.977]	9.379 [8.650, 10.108]	16.418 [15.469, 17.367]	0.429 [0.387, 0.471]
Sibling age difference					
	0 to 1 year	6.931 [6.393, 7.469]	5.292 [4.930, 5.654]	12.223 [11.650, 12.795]	0.567 [0.539, 0.595]
	2 years	6.442 [6.003, 6.880]	6.291 [5.905, 6.678]	12.733 [12.243, 13.223]	0.506 [0.479, 0.532]
	3 years	7.003 [6.404, 7.602]	5.806 [5.373, 6.240]	12.809 [12.189, 13.430]	0.547 [0.515, 0.578]
	4 years	7.229 [6.451, 8.006]	6.155 [5.605, 6.705]	13.383 [12.576, 14.191]	0.540 [0.502, 0.578]
	5 years or more	6.243 [5.654, 6.831]	6.870 [6.417, 7.323]	13.113 [12.482, 13.743]	0.476 [0.445, 0.508]
Degree of urbanization					
	Urban	6.244 [5.952, 6.537]	5.770 [5.543, 5.997]	12.015 [11.698, 12.332]	0.520 [0.503, 0.537]
	Rural	4.981 [4.584, 5.379]	6.557 [6.201, 6.913]	11.538 [11.095, 11.982]	0.432 [0.405, 0.459]

Notes: Numbers without brackets are point estimates. Numbers in brackets report 95% bootstrap confidence intervals based on 1,500 replications.

A closer look at the variance components helps clarify this difference. Among the gender-composition subsamples, sisters exhibit the largest family-specific variance component, with $\sigma^2_\alpha = 10.06$, and the smallest idiosyncratic variance component, with $\sigma^2_u = 4.94$. By contrast, σ^2_α ranges from 5.84 in mixed-gender dyads with a female firstborn to 6.60 in mixed-gender dyads with a male firstborn. These results suggest that the gender structure of educational sibling correlations in Turkey differs markedly from that documented for the developed countries studied in the literature so far. In particular, the sister-only subsample stands out because of its exceptionally large family-specific component. This pattern motivates the decomposition analysis

when we examine whether observed shared family characteristics account for the gap.

that follows, where we use our harmonic mean decomposition framework to investigate how this gender heterogeneity interacts with other observed shared family characteristics in Section 4.3.

Moreover, akin to the first-born effects discussed by Black et al. (2005), Black et al. (2011a), Black et al. (2011b), and Björklund and Jäntti (2012), our estimates reveal an interesting gender-based first-born pattern. We estimate an educational sibling correlation of 0.44 for the mixed-gender subsample in which the firstborn is female, which is statistically significantly lower than the corresponding estimate of 0.51 for the mixed-gender subsample in which the firstborn is male. According to the variance components reported in Table 2, this difference is driven by a larger family-specific variance component, σ_α^2 (6.60 vs. 5.84), and a smaller idiosyncratic variance component, σ_u^2 (6.23 vs. 7.55), in sibling dyads with male firstborns relative to those with female firstborns. This pattern directionally echoes the first-born difference in mean years of education reported in Table 1: the female sibling has 1.39 fewer years of education in the male-firstborn subsample, compared with a gap of 0.60 years in the female-firstborn subsample. Mean gaps and variance components are, however, distinct statistical objects; the coincidence in direction is suggestive rather than mechanical, and we return to the underlying drivers in Section 4.3.

Next, we examine the evolution of educational sibling correlations across birth cohorts. Focusing on the subsamples in which both siblings are born in the 1970s, 1980s, or 1990s, we observe a mild inverse U-shaped pattern: the educational sibling correlation rises slightly from 0.54 in the 1970s to 0.57 in the 1980s, before falling sharply and statistically significantly to 0.43 in the 1990s. This pattern is driven mainly by (i) the moderate increase in the family-specific variance component, σ_α^2, from 6.99 in the 1970s to 8.33 in the 1980s, and (ii) its subsequent sharp decline to 4.31 in the 1990s. As shown in Table 1, the standard deviation of years of education also follows a similar pattern, albeit to a much smaller extent, while mean years of schooling increase sharply by approximately five years over the same period. One possible explanation is the 1997 education reform, which extended compulsory schooling from five to eight years. However, because our data do not contain precise and reliable information on the exact date at which individuals started school, and because this reform affects only the tail end of our estimation sample, we refrain from drawing a definitive conclusion.[10] Consistent with the negative relationship between sibling age differences and educational sibling correlations, the mixed-decade subsamples yield lower sibling correlations on average, which is in line with their higher mean age difference between siblings (2.16 years, compared with 1.28 years in the same-decade subsamples). This pattern is also consistent with the idea that larger age gaps may reduce sibling similarity in educational outcomes, since siblings farther apart in age are more likely to experience somewhat different family circumstances, schooling environments, and macro conditions during childhood.

Our investigation of how educational sibling correlations vary with parental characteristics yields several interesting findings. First, conditioning educational sibling correlations on parental educational attainment generates a level-shift aggregation effect—a generalized form of Simpson's paradox, in which every subset-specific correlation lies below the grand correlation of 0.52, with the largest subset-specific estimate reaching only 0.47 in families where the mother has secondary education. This is a level shift rather than a directional reversal in the classical Simpson sense, and in our framework it corresponds to a large positive π under the parental-education partition.[11]

[10]Based on birth year alone, our 1990s cohort (both siblings born 1990–1999) would have been largely exposed to the eight-year compulsory schooling regime, since children born in 1990 or later were of primary school age when the reform was implemented. The 1980s cohort, by contrast, would have completed primary schooling under the previous five-year regime. The timing is therefore consistent with the sharp decline in σ_α^2 between the two cohorts, but a causal interpretation would require disentangling the reform's effect from contemporaneous trends in urbanization, female labor force participation, and overall educational expansion—all of which evolved substantially over this period. A formal evaluation of the reform's role in shaping educational sibling correlations is beyond the scope of this paper and would benefit from data with precise schooling-start information, such as that used by Cesur and Mocan (2018), Kırdar et al. (2018) and other related studies.

[11]In Section 4.2, we show that this pattern is driven by the disproportionately large between effect that arises when the sample is partitioned by parental education.

Turning first to paternal education, and setting aside the high-education category because of its imprecision due to the small number of observations, we observe a statistically meaningful inverted U-shaped pattern across the remaining three categories. In particular, the sibling correlation is 0.45 in families where the father has primary education, which is statistically significantly higher than the corresponding estimates for families in which the father has secondary education (0.38) or no education (0.37). At the same time, although the no-education and secondary-education paternal groups are not statistically distinguishable in terms of sibling correlations, they differ considerably in their underlying variance components, with the latter exhibiting markedly lower family-specific and idiosyncratic variance components.

By contrast, although the estimates by maternal education display a broadly similar pattern in the levels of sibling correlations and variance components, these differences are not statistically significant, as the corresponding confidence intervals overlap substantially across categories. Finally, somewhat surprisingly, parental age at first birth does not appear to be related to the educational sibling correlation in a statistically meaningful way.

In our analysis, educational sibling correlations are estimated using the educational attainment of the two eldest siblings in each family. At the same time, we keep track of the total number of siblings in the family and examine how family size relates to the correlation between the eldest two siblings. We find that the educational sibling correlation declines only marginally and insignificantly from 0.52 in two-child families to 0.51 in three-child families. By contrast, among families with four or more children, the educational sibling correlation falls to 0.43, which is statistically significantly lower than in the other two groups.

Partitioning the sample by the age difference between the two eldest siblings yields an interesting non-monotonic pattern. As the age difference increases from 0–1 year to 2 years, the educational sibling correlation declines significantly from 0.57 to 0.51. Increasing the age difference further to 3 and 4 years is associated with a modest rise in the sibling correlation, to 0.55 and 0.54, respectively, although these increases are not statistically significant. However, when the age difference reaches 5 years or more, the sibling correlation declines again, to 0.48.

The last family-specific variable we use to condition our estimates is the type of the current place of residence. Consistent with Öztunalı and Torul (2022), where we report a surprisingly lower intergenerational educational correlation coefficient for rural than for urban places of residence in Turkey, we find a statistically significantly lower educational sibling correlation in rural areas (0.43) than in urban areas (0.52). In Öztunalı and Torul (2022), we show that the lower rural intergenerational correlation coefficient, which may at first glance appear to indicate greater intergenerational mobility, is in fact driven by the lower probability of attaining tertiary education in rural places of residence. A similar pattern appears to be present here. In our sample, mean years of education in rural areas are 2.13 years lower than in urban areas, while the standard deviation of years of education is identical across the two groups at 3.63. This suggests that the lower rural sibling correlation should not be interpreted mechanically as evidence of greater mobility, but rather in light of the substantially lower average level of educational attainment in rural areas.

4.2 Single Layer Decomposition of Educational Sibling Correlations & Counterfactuals

We now apply the harmonic mean decomposition to identify which observed shared family characteristic contributes most through the between effect, captured by π in (15).

The results of the single-layer decomposition exercise reported in Table 3 for the REML estimates and in Table C.2 for the ANOVA-based estimates show that the grand sibling correlation is numerically very close to the weighted harmonic mean of the within-subset sibling correlations in most cases. This indicates that the between-subset amplification term is generally modest. This does not imply linear aggregation; rather, it shows that the non-linear between-subset channel is empirically small in this setting, so that when π is close to zero, the grand sibling correlation lies very near the weighted harmonic mean of the subset correlations.[12]

Table 3: Single-layer decomposition of educational sibling correlations: REML estimates

Dimension	Group	γ_k	v_k	w_k	ρ_k	V	W	$\hat{\rho}$	$\tilde{\rho}$	ρ	π
Birth cohorts	1970s	0.2688	1.8809	0.0021	0.5438	6.6901	0.0079	0.5252	0.2898	0.5247	-0.0010
	1970–1980s	0.1277	0.9419	0.0010	0.5047						
	1980s	0.2852	2.3755	0.0022	0.5675						
	1980–1990s	0.1123	0.6039	0.0009	0.5110						
	1990s	0.2057	0.8878	0.0016	0.4340						
Father's education	No education	0.1653	0.7996	0.4039	0.3734	4.2553	2.4428	0.4165	0.9586	0.5247	0.2599
	Primary	0.5450	2.7474	1.3313	0.4518						
	Secondary	0.2242	0.6440	0.5479	0.3765						
	Higher	0.0653	0.0642	0.1597	0.2259						
Mother's education	No education	0.4408	2.4416	0.8405	0.4304	4.7913	1.9067	0.4446	0.9581	0.5247	0.1800
	Primary	0.4583	2.1245	0.8739	0.4637						
	Secondary	0.0834	0.2183	0.1592	0.4730						
	Higher	0.0173	0.0069	0.0331	0.1157						
First fatherhood age	Below 20s	0.0345	0.1964	0.0103	0.4464	6.3994	0.2985	0.5128	1.0418	0.5247	0.0232
	20s	0.6345	3.9003	0.1894	0.5183						
	30s	0.2676	1.8683	0.0799	0.5150						
	40s or above	0.0632	0.4344	0.0189	0.4900						
First motherhood age	Below 20s	0.2127	1.2396	0.0499	0.5025	6.4632	0.2348	0.5167	0.9111	0.5247	0.0154
	20s	0.6352	4.1723	0.1491	0.5260						
	30s	0.1412	0.9893	0.0332	0.5042						
	40s or above	0.0107	0.0620	0.0025	0.4219						
Sibling gender composition	Brothers	0.3674	2.2184	-0.0472	0.5431	6.8265	-0.1286	0.5386	-1.4134	0.5247	-0.0258
	Sisters	0.1692	1.7036	-0.0218	0.6710						
	Male/Female	0.2650	1.7470	-0.0341	0.5141						
	Female/Male	0.1981	1.1575	-0.0255	0.4363						
Sibling count	Two	0.6515	3.6520	0.3471	0.5217	6.1653	0.5327	0.5046	0.9715	0.5247	0.0397
	Three	0.2422	1.7662	0.1291	0.5083						
	Four or more	0.1061	0.7470	0.0566	0.4287						
Sibling age difference	0 to 1 year	0.2137	1.4811	-0.0015	0.5670	6.7051	-0.0072	0.5246	0.4329	0.5247	0.0002
	2 years	0.2981	1.9203	-0.0021	0.5059						
	3 years	0.1838	1.2871	-0.0013	0.5467						
	4 years	0.1197	0.8652	-0.0009	0.5401						
	5 years or more	0.1844	1.1514	-0.0013	0.4761						
Degree of urbanization	Urban	0.6761	4.2218	0.5836	0.5197	5.8349	0.8631	0.4920	0.9523	0.5247	0.0664
	Rural	0.3238	1.6131	0.2795	0.4317						

Notes: The table reports the single-layer decomposition results based on the REML estimates obtained following Mazumder (2008). V, W, $\hat{\rho}$, $\tilde{\rho}$, ρ, and π are partition-level quantities and are therefore shown once for each partition.

Both tables also illustrate the trade-off between $\hat{\rho}$ and π across partitioning schemes. When a partitioning scheme produces subsets that are more similar to one another, in the sense that the subset-specific means $\overline{\alpha}_k$ and $\overline{u}_k$ lie closer to their corresponding grand means $\overline{\alpha}$ and $\overline{u}$, the between effect π declines and the weighted harmonic mean of subset sibling correlations, $\hat{\rho}$, moves closer to the grand sibling correlation, ρ. At the same time, a large $\tilde{\rho}$ relative to $\hat{\rho}$ does not, by itself, guarantee a sizeable π, because the magnitude of the between effect also depends on the variance-based weights V and W in (14), which are highly sensitive to the choice of partitioning scheme. For example, the birth-cohort partitioning in Table 3 yields a between correlation of only $\tilde{\rho} = 0.29$, well below the grand correlation $\rho = 0.52$. Yet because its weight is very small, $W = 0.0079$ relative to $V = 6.6901$, the downward pull exerted by the between correlation is not strong enough to move the grand correlation materially away from $\hat{\rho}$.

Table 3 also surfaces a feature of REML-based estimates that becomes visible only through our decomposition. In the sibling-gender and sibling-age-difference partitions, the decomposition yields negative between-subset weights w_k, and in the gender-based partition it produces a between correlation, $\tilde{\rho} = -1.41$, that lies outside the admissible range for a variance ratio. This arises not from a flaw in the decomposition itself but from the fact that REML estimates the grand and subset-specific variance components through sepa-

[12]Throughout our decomposition analysis, we report point estimates rather than confidence intervals for the decomposition objects. Because these objects are deterministic functions of the underlying variance components, whose uncertainty is quantified in Table 2, their uncertainty is directly inherited from the uncertainty in the underlying estimates.

rate unconstrained optimizations, with no cross-equation restriction enforcing $\sigma_\alpha^2 \geq \sum_{k=1}^{K} \gamma_k \sigma_{\alpha,k}^2$. When this inequality is violated in finite samples—as happens here for the gender and age-difference partitions—the implied between component, recovered residually via (7), turns negative. Our framework makes this internal inconsistency visible because it requires the grand and subset-level variance components to be mutually coherent at the coefficient level; standard variance-component analyses, which examine σ_α^2 and $\sigma_{\alpha,k}^2$ in isolation, do not.

Importantly, this artifact does not affect the substantive interpretation of the gender-partition results. The negative $\tilde{\rho}$ and w_k terms reflect the small magnitude of the between-subset variation across gender profiles relative to the grand variance, which generates a near-zero π regardless of how the underlying variance components are estimated. The corresponding ANOVA-based decomposition reported in Table C.2, which by construction satisfies the additive consistency between grand and subset-level variance components, yields $\tilde{\rho} = 0.945$, $\hat{\rho} = 0.520$, and $\pi = 0.001$ for the same partition, confirming that the between effect of the gender partition is empirically negligible. The REML- and ANOVA-based decompositions are otherwise extremely similar across all partitioning schemes, and our main conclusions are unaffected by the choice of estimator.

Overall, the findings indicate that although the between correlation $\tilde{\rho}$ is generally larger than the weighted harmonic mean of subset correlations $\hat{\rho}$, its influence on the grand sibling correlation is typically limited by the small weight attached to the between component in the harmonic aggregation formula. As a result, the grand educational sibling correlation remains close to $\hat{\rho}$ for most partitioning schemes based on the observed shared family characteristics available in our data. Most importantly, regardless of whether we use REML- or ANOVA-based estimates, the decomposition results consistently show that partitioning the sample by paternal and maternal education yields the largest between effects. These schemes produce the largest π values in both sets of estimates: 0.26 and 0.18 under REML, and 0.25 and 0.17 under ANOVA, for paternal and maternal education, respectively.

Next, to explore the extent to which subsets shape the grand educational sibling correlation ρ, we perform subset-removal counterfactuals by (i) calculating the grand correlation that would arise upon the removal of the m^{th} subset while holding the between-correlation constant, which quantifies the change in ρ through the change in $\hat{\rho}$ among the remaining $K-1$ subsets, and (ii) estimating the actual grand sibling correlation that results from removing this subset, which reflects the joint change in $\hat{\rho}$ and $\tilde{\rho}$. To keep the discussion focused, we report the subset-removal counterfactuals only for the partitioning schemes that yield the largest π values, namely paternal and maternal education. Since π summarizes the strength of the between effect in our framework, these are also the cases in which the counterfactual decomposition is most informative.[13]

The results of the counterfactual exercise summarized in Table 4 yield an interesting finding: removing the subset with the largest within-subset correlation—families in which the father has primary educational attainment—increases the grand sibling correlation by 0.062, revealing another Simpson-type feature of the aggregation process. Moreover, most of this change operates through the within-channel, as the corresponding $\dot{\Delta}_m$ term is 0.060, whereas the contribution of the between-channel, captured by $\ddot{\Delta}_m$, is negligible at only 0.003. These results suggest that even when a partitioning scheme generates a sizeable between effect, as reflected in a large π term, the impact of removing individual subsets within that partition need not operate primarily through changes in the between-subset structure. In the paternal- and maternal-education partitions studied here, the between structure among the remaining subsets remains remarkably stable after removal, so the counterfactual change in the grand correlation arises almost entirely through the reshaping of the harmonic mean of the remaining subset correlations.

[13] For brevity, we confine the discussion to the paternal- and maternal-education partitioning schemes, which yield the largest π values. The corresponding counterfactual results for the remaining partitioning schemes are available upon request.

Table 4: Subset-removal counterfactuals for paternal and maternal education (REML estimates)

Excluded subset	ρ^c_{-m}	π^c_{-m}	$\dot{\Delta}_m$	$\ddot{\Delta}_m$	Δ_m
Panel A: Father's education					
No education	0.501	0.171	-0.023	-0.002	-0.026
Primary	0.584	0.603	0.060	0.003	0.062
Secondary	0.513	0.208	-0.012	-0.001	-0.012
Higher	0.501	0.187	-0.024	0.001	-0.023
Panel B: Mother's education					
No education	0.499	0.084	-0.026	-0.002	-0.028
Primary	0.536	0.246	0.012	-0.001	0.011
Secondary	0.500	0.128	-0.025	0.000	-0.024
Higher	0.518	0.161	-0.006	0.000	-0.006

Notes: ρ^c_{-m} denotes the counterfactual grand sibling correlation after removing subset m while holding the between-subset correlation fixed. π^c_{-m} is the associated counterfactual Simpson's paradox coefficient. $\dot{\Delta}_m$ and $\ddot{\Delta}_m$ denote the within- and between-channel components of the absolute change in the grand sibling correlation, respectively, and Δ_m is their sum. Positive values of Δ_m indicate that removing the subset raises the grand sibling correlation, whereas negative values indicate that it lowers it. See Table C.3 and Table C.4 for the full counterfactual components based on REML and ANOVA, respectively.

Table C.3 provides a detailed account of the mechanism underlying this surprising result. Removing the paternal primary-education subset reduces V, the weight attached to $\hat{\rho}$ in the harmonic formulation of ρ, from 4.25 to 3.31, while increasing W, the weight attached to $\tilde{\rho}$, from 2.44 to 5.11. At the same time, $\hat{\rho}$ declines from 0.42 to 0.36, whereas $\tilde{\rho}$ rises only slightly from 0.96 to 0.97. Hence, the increase in ρ is driven only to a very limited extent by changes in the levels of the within- and between-subset correlations themselves. Rather, it is driven primarily by the sharp increase in the relative weight of the between correlation, which is already substantially larger than the within correlation, in the weighted harmonic mean representation. In other words, removing this subset changes the geometry of the aggregation much more than it changes the underlying within- and between-correlation levels. A similar mechanism is at work when we examine the decline in ρ following the removal of the no-education, secondary-education, and higher-education paternal subsets: although $\hat{\rho}$ and $\tilde{\rho}$ change only minimally, these removals lower ρ by shifting the V/W ratio in favor of V, thereby reducing the relative weight of the between correlation in the harmonic aggregation. This again shows that the subset-removal counterfactuals are driven mainly by reweighting within the harmonic aggregation formula rather than by large changes in the within- and between-subset correlation levels themselves. Thus, the key source of counterfactual variation is not necessarily a large movement in the local and between correlation objects themselves, but a change in the weight structure through which they are aggregated.

The mechanism uncovered here is not obvious from conventional variance decompositions. Because those approaches operate additively on the variance component underlying the sibling correlation, they do not reveal the nonlinear aggregation structure of the correlation coefficient itself. As a result, they cannot directly uncover the reweighting mechanism through which relatively stable within- and between-correlation objects generate counterfactual changes in the grand sibling correlation.

4.3 Dual Layer Decomposition of Educational Sibling Correlations

In the last stage of our analysis, exploiting the recursive nature of our harmonic mean decomposition approach, we explore the drivers of (i) the high educational sibling correlation exhibited by the sisters subsample, and (ii) the gender-based firstborn effect first discussed in Section 4.1. Table 5 summarizes how the observed shared familial characteristics we have explored in our study interact with the sisters vs. brothers difference and the gender-specific firstborn effect our prior results revealed while Table C.5 gives a full-account of our dual-layer decomposition exercise.

Table 5: Dual-layer decomposition of gender-based gaps in educational sibling correlations (REML estimates)

Second-layer partition	Within-subsubset effect	Between-subsubset effect	Total effect
Panel A: Sisters vs. Brothers			
Birth cohorts	0.130	-0.002	0.128
Father's education	0.142	-0.014	0.128
Mother's education	0.110	0.018	0.128
First fatherhood age	0.125	0.002	0.128
First motherhood age	0.122	0.006	0.128
Sibling count	0.137	-0.009	0.128
Sibling age difference	0.128	-0.001	0.128
Degree of urbanization	0.117	0.011	0.128
Panel B: Female-firstborn vs. Male-firstborn mixed-gender dyads			
Birth cohorts	-0.078	0.000	-0.078
Father's education	-0.107	0.029	-0.078
Mother's education	-0.091	0.014	-0.078
First fatherhood age	-0.077	0.000	-0.078
First motherhood age	-0.094	0.016	-0.078
Sibling count	-0.085	0.007	-0.078
Sibling age difference	-0.085	0.007	-0.078
Degree of urbanization	-0.090	0.012	-0.078

Notes: Panel A decomposes the gap in educational sibling correlations between sisters and brothers. Panel B decomposes the gap between mixed-gender dyads with a female firstborn and mixed-gender dyads with a male firstborn. The first term in (19) is labeled the within-subsubset effect and the second term the between-subsubset effect. Corresponding ANOVA-based results are extremely similar and are omitted for brevity.

Our findings indicate that the *between-subsubset effect*, which captures the role of heterogeneity induced by the second layer of the decomposition based on observed shared familial characteristics, explains only a negligible part of the sisters-versus-brothers difference in educational sibling correlations. By contrast, an overwhelmingly large part of this difference is accounted for by the *within-subsubset effect*, which reflects the change in the harmonic mean of the subsubset correlations between the two gender subsets. Furthermore, we find that the gender-based firstborn effect exhibits a similar pattern and is likewise driven primarily by a disproportionately large *within-subsubset effect*. These results suggest that the sisters-versus-brothers gap and the gender-based firstborn effect are not primarily driven by the heterogeneities induced by the observed shared family characteristics considered here. Because the within-subsubset effect dominates and the between-subsubset effect remains small across all second-layer partitioning schemes, these gender-based differences appear to persist within the observed categories themselves.

Four candidate mechanisms, all consistent with the Turkish institutional context, could generate the pattern we observe. First, and most directly, families may differ in their elasticity of educational investment with respect to daughters' relative to sons'. If a son's schooling is treated as a near-universal normative target while a daughter's schooling is contingent on family-specific attitudes, resources, or constraints, the family-specific variance component σ^2_α will mechanically be larger for sisters than for brothers, which is precisely what Table 2 shows. Second, marriage-market considerations are known to shape the timing of daughters' withdrawal from education in Turkey and in comparable settings (Kırdar et al., 2018), and because marriage-market opportunities are correlated within families, they would induce sibling similarity for sisters without leaving a corresponding mark on brothers. Third, intra-household bargaining models imply that daughters' schooling is more responsive than sons' to the relative position of the mother within the household, a mechanism for which Rankin and Aytaç (2006) and Smits and Hoşgör (2006) provide direct Turkish evidence. Fourth, local community-level gender norms—conservative versus liberal—affect daughters' schooling more sharply than sons', and because siblings share a neighborhood and a community, these norms load onto the family component rather than the idiosyncratic component. The coarse rural/urban partition in our data does not capture this fine-grained cultural variation. A common feature of all four channels is that they operate through family- or community-level unobservables that are not captured by the categorical observables in our dataset, which is consistent with the dominance of the within-subsubset effect documented above. Discriminating between these mechanisms requires data with richer measures of intra-household allocation, marriage-market exposure, and community gender norms, and constitutes a natural extension.

5 Conclusion

This paper develops a decomposition framework for sibling correlations that operates directly on the correlation coefficient rather than on its variance components. The grand sibling correlation is exactly a weighted harmonic mean of within- and between-subset correlations, with variance-based weights that make the relative importance of each channel transparent. The framework is recursive, allowing progressive refinement of the partitioning structure while preserving the exact relationship to the population-wide measure, and supports a counterfactual subset-removal exercise that separates each subset's contribution into a harmonic-mean channel and a between-subset channel. Relative to existing variance-component decompositions, our approach makes the non-linear aggregation structure of the sibling correlation explicit and shows how subgroup partitions shape the population-wide measure through channels that additive decompositions cannot, by construction, uncover.

Applying this framework to Turkish Demographic and Health Survey data, we obtain three findings that contribute to the empirical literature on sibling correlations. First, we document an overall educational sibling correlation of approximately 0.52 for Turkey, which—bearing in mind the data-comparability caveats discussed in Section 3.1—places the country broadly above the Nordic benchmark but below the United States and Germany. Partitioning the sample by parental education generates the largest between-subset amplification among all observable dimensions we examine, producing a level-shift aggregation effect—a generalized Simpson's paradox—in which every subset-specific correlation lies below the grand correlation. Second, the sister-only subsample exhibits an educational sibling correlation 0.13 points higher than the brother-only estimate—a gap substantially larger than those typically reported for developed countries and, in the United States and Germany, opposite in sign. Applying our decomposition recursively, we find that this gap, together with a related gender-based firstborn effect, persists within every observable partition of family characteristics considered, suggesting that unobserved gender-specific mechanisms—plausibly related to gendered parental investment attitudes, intra-household bargaining, or marriage-market considerations—play a central role in shaping the Turkish pattern. Third, our counterfactual exercise reveals that the subset-removal effects on the grand sibling correlation operate primarily through reweighting within the harmonic aggregation formula rather than through changes in the underlying within- and between-subset correlation levels. This reweighting mechanism is a structural feature of the harmonic aggregation that additive variance decompositions cannot, by construction, detect.

Several directions for future research emerge from our analysis. First, extending the framework to outcomes beyond education—earnings, occupational status, or health—would allow researchers to assess whether the aggregation patterns documented here are specific to educational attainment or reflect more general features of how family background shapes socioeconomic outcomes. Second, a formal investigation of the gender-specific mechanisms underlying the sister-brother and firstborn gaps we document would require data with richer measures of intra-household allocation, gender attitudes, and marriage-market exposure than the DHS provides. Third, a direct numerical comparison between our coefficient-level decomposition and the variance-component decompositions of Karlson and In (2025) and Karlson (2025) would further clarify the relationship between the two frameworks and the conditions under which each is most informative. Fourth, because the decomposition requires only a sibling sample and subset indicators, it ports naturally to other DHS-covered low- and middle-income countries, and should be a useful building block for extending the comparative mobility literature to settings where linked parent–child administrative data are unavailable.

For inequality policy, two implications follow directly. Among the observable dimensions we examine, parental education generates the largest between-subset amplification of the educational sibling correlation, which identifies it as the stratifying margin where a given reduction in between-group heterogeneity would produce the largest reduction in the population-wide measure of family-background inequality. The persistence of the sister-brother gap within every observable partition, by contrast, indicates that the stronger relative role of shared family background in sisters' schooling is not reducible to the usual observable characteristics, so closing this gap requires instruments that target the unobservables—gendered parental investment, intra-household allocation, community norms—rather than standard socioeconomic margins. A methodological caveat is worth stating explicitly: our decomposition identifies the structure and location of these gaps but does not, on its own, identify the causal mechanisms that generate them; that task requires data with richer measures of the unobservables and is left to future work.

More broadly, the central message of the paper is that sibling correlations aggregate harmonically, not additively, and that the distinction matters: the channels through which family background shapes the population-wide measure are not the same as the channels through which it shapes its variance components. Treating the correlation coefficient as the primitive object, rather than as a derived ratio, brings these channels into view.

Author contributions. Both authors contributed equally to all aspects of this work, including conceptualization, methodology, formal analysis, software, data curation, writing, and review.

References

Ahsan, M. N., Emran, M. S., Jiang, H., Han, Q., and Shilpi, F. (2023). Growing up together: Sibling correlation, parental influence, and intergenerational educational mobility in developing countries. Policy Research Working Paper 10285, World Bank.

Ahsan, M. N., Emran, M. S., Jiang, H., and Shilpi, F. (2025). Making the most of coresident data: Credible evidence on intergenerational mobility with sibling correlation. Journal of Development Economics, 176:103508.

Anger, S. and Schnitzlein, D. D. (2017). Cognitive skills, non-cognitive skills, and family background: Evidence from sibling correlations. Journal of Population Economics, 30(2):591–620.

Bingley, P. and Cappellari, L. (2019). Correlation of brothers' earnings and intergenerational transmission. The Review of Economics and Statistics, 101(2):370–383.

Björklund, A. and Jäntti, M. (2009). Intergenerational income mobility and the role of family background. In Salverda, W., Nolan, B., and Smeeding, T. M., editors, The Oxford Handbook of Economic Inequality, pages 491–521. Oxford University Press.

Björklund, A. and Jäntti, M. (2012). How important is family background for labor-economic outcomes? Labour Economics, 19(4):465–474.

Björklund, A. and Jäntti, M. (2020). Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. Research in Social Stratification and Mobility, 70:100455.

Björklund, A., Lindahl, L., and Lindquist, M. J. (2010). What more than parental income, education and occupation? an exploration of what swedish siblings get from their parents. The B.E. Journal of Economic Analysis & Policy, 10(1).

Björklund, A. and Salvanes, K. G. (2011). Education and family background: Mechanisms and policies. In Handbook of the Economics of Education, volume 3, pages 201–247. Elsevier.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. American Economic Review, 95(1):437–449.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011a). Older and wiser? birth order and IQ of young men. CESifo Economic Studies, 57(1):103–120.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011b). Too young to leave the nest? the effects of school starting age. The Review of Economics and Statistics, 93(2):455–467.

Bredtmann, J. and Smith, N. (2018). Inequalities in educational outcomes: How important is the family? Oxford Bulletin of Economics and Statistics, 80(6):1117–1144.

Cesur, R. and Mocan, N. (2018). Education, religion, and voter preference in a muslim country. Journal of Population Economics, 31(1):1–44.

Dayioğlu, M. (2005). Patterns of change in child labour and schooling in turkey: The impact of compulsory schooling. Oxford Development Studies, 33(2):195–210.

Fletcher, J. M., Lu, Q., Mazumder, B., and Song, J. (2023). Understanding sibling correlations in education: Molecular genetics and family background. Iza discussion paper, IZA.

Forsberg, E., Khan, A., and Rosenqvist, O. (2025). Do sibling correlations in skills, schooling, and earnings vary by socioeconomic background? insights from sweden. Labour Economics, 96:102741.

Hällsten, M. (2026). Split sibling correlations aren't what you think: The hidden mean differences. Manuscript, Stockholm University.

Hertz, T. (2008). A group-specific measure of intergenerational persistence. Economics Letters, 100(3):415–417.

Karlson, K. B. (2025). What do group-specific sibling correlations really measure? SocArXiv. SocArXiv preprint.

Karlson, K. B. and In, J. (2025). Decomposing sibling correlations: A new measure of group-specific intergenerational persistence in socioeconomic outcomes. European Sociological Review, 41(5):822–830.

Kırdar, M. G., Dayıoğlu, M., and Koç, İ. (2018). The effects of compulsory-schooling laws on teenage marriage and births in turkey. Journal of Human Capital, 12(4):640–668.

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. Journal of Population Economics, 21(3):685–701.

Muñoz, E. and Jaque, G. (2026). Sibling correlations in schooling around the world: A new database. World Development, 204:107415.

Öztunalı, O. and Torul, O. (2022). The evolution of intergenerational educational mobility in turkey. Emerging Markets Finance and Trade, 58(14):4033–4049.

Rankin, B. H. and Aytaç, I. A. (2006). Gender inequality in schooling: The case of turkey. Sociology of Education, 79(1):25–43.

Schnitzlein, D. D. (2014). How important is the family? evidence from sibling correlations in permanent earnings in the USA, germany, and denmark. Journal of Population Economics, 27(1):69–89.

Smits, J. and Hoşgör, A. G. (2006). Effects of family background characteristics on educational participation in turkey. International Journal of Educational Development, 26(5):545–560.

Solon, G. (1999). Intergenerational mobility in the labor market. In Ashenfelter, O. C. and Card, D., editors, Handbook of Labor Economics, volume 3, chapter 29, pages 1761–1800. Elsevier.

Solon, G., Corcoran, M., Gordon, R., and Laren, D. (1991). A longitudinal analysis of sibling correlations in economic status. The Journal of Human Resources, 26(3):509–534.

Appendix

A. Decomposition of the Educational Sibling Correlation

Start with the sample-wide variance of the family-specific random component σ^2_α:

$$\sigma^2_\alpha = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(\alpha_i - \overline{\alpha})^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i} 1 - 1} \tag{A.1}$$

where N stands for the total number of families in the sample, J_i refers to the number of siblings in family i, and $\overline{\alpha}$ is the grand mean of the impact of family-specific unobservables. Now, decompose the sample into K mutually exclusive subsets and re-express the variance as a cross-sectional sum across subsets:

$$\sigma^2_\alpha = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{J_{i,k}}(\alpha_{i,k} - \overline{\alpha})^2}{(\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}) - 1} \tag{A.2}$$

where $\alpha_{i,k}$ stands for the impact of family-specific unobservables and $J_{i,k}$ corresponds to the number of siblings in family i of subset k while N_k denotes the total number of families in subset k.

Now focus on the numerator of (A.2) and add and subtract the k[th] subset's mean effect via unobservable family-specific traits $\overline{\alpha}_k$ in the quadratic term in parentheses:

$$\begin{aligned}
\sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{J_{i,k}}(\alpha_{i,k} - \overline{\alpha})^2 &= \sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha})^2 = \sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha} \pm \overline{\alpha}_k)^2 \\
&= \sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}[(\alpha_{i,k} - \overline{\alpha}_k)^2 + (\overline{\alpha}_k - \overline{\alpha})^2 + 2(\overline{\alpha}_{i,k} - \overline{\alpha}_k)(\overline{\alpha}_k - \overline{\alpha})] \\
&= \sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}[(\alpha_{i,k} - \overline{\alpha}_k)^2 + (\overline{\alpha}_k - \overline{\alpha})^2] + \underbrace{2\sum_{k=1}^{K}(\overline{\alpha}_k - \overline{\alpha})\left[\sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha}_k)\right]}_{*}
\end{aligned} \tag{A.3}$$

Note that the last term in (A.3) is zero due to the fact that $\sum_{i=1}^{N_k} J_{i,k}\alpha_{i,k} = \sum_{i=1}^{N_k} \overline{\alpha}_k$. Therefore,

$$\sigma^2_\alpha = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k}\sum_{j=1}^{J_{i,k}}(\alpha_{i,k} - \overline{\alpha})^2}{(\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}) - 1} = \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}[(\alpha_{i,k} - \overline{\alpha}_k)^2 + (\overline{\alpha}_k - \overline{\alpha})^2]}{(\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}) - 1} \tag{A.4}$$

Next, define the variance of α in subset k:

$$\sigma^2_{\alpha,k} = \frac{\sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha}_k)^2}{(\sum_{i=1}^{N_k} J_{i,k}) - 1} \tag{A.5}$$

Denote the number of siblings in subset k and the entire sample with $n(k) = \sum_{i=1}^{N} J_{i,k}$ and $\sum_{k=1}^{K} n(k) = S$, respectively. Let $\gamma_k = \frac{n(k)-1}{S-1}$ and $\eta_k = \frac{n(k)}{S-1}$ be two subset-specific weight terms. Note that the numerator of (A.4) has a component that is in additive form with respect to $\sigma^2_{\alpha,k}[n(k)-1] = \sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha}_k)^2$:

$$\begin{aligned}
\sigma^2_\alpha &= \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}(\alpha_{i,k} - \overline{\alpha}_k)^2}{(\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}) - 1} + \frac{\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}(\overline{\alpha}_k - \overline{\alpha})^2}{(\sum_{k=1}^{K}\sum_{i=1}^{N_k} J_{i,k}) - 1} \\
&= \frac{\sum_{k=1}^{K}[n(k)-1]\sigma^2_{\alpha,k}}{S-1} + \frac{\sum_{k=1}^{K} n(k)(\overline{\alpha}_k - \overline{\alpha})^2}{S-1} \\
&= \underbrace{\sum_{k=1}^{K} \gamma_k \sigma^2_{\alpha,k}}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}_{\substack{\text{between} \\ \text{component}}}
\end{aligned} \tag{A.6}$$

According to (A.6), the grand variance of α is equal to the sum of (i) subset-size weighted average of subset variances ($\sum_{k=1}^{K} \gamma_k \sigma^2_{\alpha,k}$ - i.e., the *within component*), and (ii) a subset-size weighted of the deviations of the subset mean from the grand mean ($\sum_{k=1}^{K} \eta_k (\overline{\alpha}_k - \overline{\alpha})^2$ - that is, the *between component*). While the grand variance σ^2_α and the *within component* can be retrieved via REML or ANOVA, the between component can only be retrieved residually - that is, after estimating the grand and subset-specific variances, one can obtain the between component by subtracting the within component from the grand variance of α. Following this cross-sectional variance decomposition logic, one can decompose the grand variances of the composite error term σ^2_ϵ and the idiosyncratic term σ^2_u in the same manner:

$$\sigma^2_u = \underbrace{\sum_{k=1}^{K} \gamma_k \sigma^2_{u,k}}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{u}_k - \overline{u})^2}_{\substack{\text{between} \\ \text{component}}} \tag{A.7}$$

$$\sigma^2_\epsilon = \underbrace{\sum_{k=1}^{K} \gamma_k \sigma^2_{\epsilon,k}}_{\substack{\text{within} \\ \text{component}}} + \underbrace{\sum_{k=1}^{K} \eta_k (\overline{\epsilon}_k - \overline{\epsilon})^2}_{\substack{\text{between} \\ \text{component}}} \tag{A.8}$$

Note that the *between components* are *variance-like* objects in the sense that are the sample-size weighted sums of sample means' deviations from the grand mean. These are truly between objects as they utilize only information at the subset mean and grand mean levels instead of individual level information. However, they are *variance-like* objects because unlike a true between-subset variance where the deviation of each subset's mean from the grand mean is only used once and the cross-sectional sum of these quadratic objects is subjected to a $K-1$ degrees of freedom correction, here subset means' deviation from the grand mean are weighted with subset sizes and the entire sum is then normalized with $S-1$. Because of their variance-like structure, we start addressing these terms with $\tilde{\sigma}^2$ from now on:

$$\tilde{\sigma}^2_\alpha = \sum_{k=1}^{K} \eta_k (\overline{\alpha}_k - \overline{\alpha})^2 \tag{A.9}$$

$$\tilde{\sigma}^2_u = \sum_{k=1}^{K} \eta_k (\overline{u}_k - \overline{u})^2 \tag{A.10}$$

$$\tilde{\sigma}^2_\epsilon = \sum_{k=1}^{K} \eta_k (\overline{\epsilon}_k - \overline{\epsilon})^2 \tag{A.11}$$

Next, we turn to the decomposition of the grand educational sibling correlation ρ:

$$\rho = \frac{\sigma_\alpha^2}{\sigma_\epsilon^2} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} \tag{A.12}$$

Let ρ_k be the educational sibling correlation of the k$^{\text{th}}$ subset:

$$\rho_k = \frac{\sigma_{\alpha,k}^2}{\sigma_{\epsilon,k}^2} = \frac{\sigma_{\alpha,k}^2}{\sigma_{\alpha,k}^2 + \sigma_{u,k}^2} \tag{A.13}$$

And let $\tilde{\rho}$ stand for the between-subset analogue of the educational sibling correlation based on the variance-like objects:

$$\tilde{\rho} = \frac{\tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\epsilon^2} = \frac{\tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\alpha^2 + \tilde{\sigma}_u^2} = \frac{\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2 + \sum_{k=1}^K \eta_k (\overline{u}_k - \overline{u})^2} \tag{A.14}$$

Then, express ρ in terms of the weighted sums of subset-specific variances and between-subset variance-like objects via (A.6), (A.7), (A.8), (A.9), (A.10), and (A.11):

$$\rho = \frac{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2 + \sum_{k=1}^K \gamma_k \sigma_{u,k}^2 + \sum_{k=1}^K \eta_k (\overline{u}_k - \overline{u})^2} \tag{A.15}$$

Assuming that partitioning yields non-zero $\sigma_{\alpha,k}^2$ and $\sigma_{u,k}^2$ terms $\forall k \in \{1, \ldots, K\}$, and non-zero variance-like between objects $\tilde{\sigma}_\alpha^2$ and $\tilde{\sigma}_u^2$, divide and multiply the subset specific variance terms in the denominator of (A.15) with $\sigma_{\alpha,k}^2$ while dividing and multiplying between subset objects with $\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2$:

$$\begin{aligned}
\rho &= \frac{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \gamma_k \sigma_{u,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2 + \sum_{k=1}^K \eta_k (\overline{u}_k - \overline{u})^2} \\
&= \frac{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \gamma_k [\sigma_{\alpha,k}^2 + \sigma_{u,k}^2] + \sum_{k=1}^K \eta_k [(\overline{\alpha}_k - \overline{\alpha})^2 + (\overline{u}_k - \overline{u})^2]} \\
&= \frac{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \gamma_k [\sigma_{\alpha,k}^2 + \sigma_{u,k}^2] \frac{\sigma_{\alpha,k}^2}{\sigma_{\alpha,k}^2} + \sum_{k=1}^K \eta_k [(\overline{\alpha}_k - \overline{\alpha})^2 + (\overline{u}_k - \overline{u})^2] \frac{\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}} \\
&= \frac{\sum_{k=1}^K \gamma_k \sigma_{\alpha,k}^2 + \sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\sum_{k=1}^K \gamma_k \frac{\sigma_{\alpha,k}^2}{\rho_k} + \frac{\sum_{k=1}^K \eta_k (\overline{\alpha}_k - \overline{\alpha})^2}{\tilde{\rho}}}
\end{aligned} \tag{A.16}$$

Now, let $v_k = \gamma_k \sigma_{\alpha,k}^2$, $\sum_{k=1}^K v_k = V$, $w_k = \eta_k (\overline{\alpha}_k - \overline{\alpha})^2$, and $W = \sum_{k=1}^K w_k = \tilde{\sigma}_\alpha^2$. Then, (A.16) can be expressed as:

$$\rho = \frac{V + W}{\sum_{k=1}^K \frac{v_k}{\rho_k} + \frac{W}{\tilde{\rho}}} \tag{A.17}$$

Note that the term $\sum_{k=1}^K \frac{v_k}{\rho_k}$ in the denominator of (A.17) is equal to $\frac{\sum_{k=1}^K v_k}{\hat{\rho}} = \frac{V}{\hat{\rho}}$ where $\hat{\rho}$ is the weighted harmonic mean of the subset-specific educational sibling correlations ρ_k with the weights v_k. Taking this into account, the grand educational sibling correlation ρ can be expressed as a weighted harmonic mean of (i) the weighted harmonic mean of within-subset sibling correlations ($\hat{\rho}$) and (ii) the between-subset sibling correlation $\tilde{\rho}$ with the subset-size weighted variance based weights V and W, respectively:

$$\rho = \frac{V + W}{\frac{V}{\hat{\rho}} + \frac{W}{\tilde{\rho}}} \tag{A.18}$$

B. Formulas Adapted from Solon et al. (1991) to be Used with Cross-Sectional Data

To estimate the educational sibling correlation according to Solon et al. (1991), one first needs to estimate the regression equation specified in equation (1) with OLS and retrieve the residuals of this regression in order to obtain the demeaned years of education $\tilde{Y}_{ij} = \epsilon_{ij}$. Next, by applying the following two ANOVA formulas adapted from Solon et al. (1991) to $\tilde{Y}_{ij}$ obtained in the previous step, one can calculate the educational sibling correlation ρ:[14]

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(\epsilon_{ij} - \bar{\epsilon}_i)^2}{S - N} \tag{B.1}$$

$$\sigma_\alpha^2 = \frac{\sum_{i=1}^{N} J_i(\bar{\epsilon}_i - \bar{\epsilon})^2}{S - 1} - \frac{\sigma_u^2(N-1)}{S-1} \tag{B.2}$$

where J_i refers to the number of siblings in family i, N stands for the total number of families, and S corresponds to the sample size.

Start the procedure by estimating the composite error term ϵ_i by regressing individuals' educational attainment Y on a set of controls X (which consists of a series of birth year and survey wave fixed effects in our case) and estimating the following regression equation with OLS:

$$Y_{ij} = \Omega + \phi X_{ij} + \epsilon_{ij} \tag{B.3}$$

where the composite error term ϵ_{ij} is the sum of family-specific random effect α_i and individual-specific random error term u_i:

$$\epsilon_{ij} = \alpha_i + u_{ij} \tag{B.4}$$

Define the grand and family-specific means of the composite error terms, i.e., $\bar{\epsilon}$ and $\bar{\epsilon}_i$, in the following way:

$$\bar{\epsilon} \equiv \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\epsilon_{ij}}{S} \tag{B.5}$$

$$\bar{\epsilon}_i \equiv \frac{\sum_{j=1}^{J_i}\epsilon_{ij}}{J_i} \tag{B.6}$$

As a result:

$$\bar{\epsilon} \equiv \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\epsilon_{ij}}{S} = \frac{\sum_{i=1}^{N} J_i\bar{\epsilon}_i}{S} \tag{B.7}$$

Now, focus on the variance of the idiosyncratic error term σ_u^2:

[14]The ANOVA formulas described in Solon et al. (1991) cannot be directly incorporated to our analysis since the authors of that study model the sibling correlations in earnings where siblings' earnings are observed at different times and ages. Thus, they need to incorporate a random time coefficient in (1) - which changes the strategy one needs to use to retrieve the sibling correlation.

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u})^2}{\sum_{i=1}^{N} J_i - 1} \tag{B.8}$$

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i+\overline{u}_i-\overline{u})^2}{\sum_{i=1}^{N} J_i - 1}$$

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{\sum_{i=1}^{N} J_i - 1} + 2\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)(\overline{u}_i-\overline{u})}{\sum_{i=1}^{N} J_i - 1} + \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(\overline{u}_i-\overline{u})^2}{\sum_{i=1}^{N} J_i - 1}$$

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{\sum_{i=1}^{N} J_i - 1} + 2\underbrace{\frac{\sum_{i=1}^{N}(\overline{u}_i-\overline{u})\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)}{\sum_{i=1}^{N} J_i - 1}}_{*} + \frac{\sum_{i=1}^{N} J_i(\overline{u}_i-\overline{u})^2}{\sum_{i=1}^{N} J_i - 1}$$

where u_{ij} is the idiosyncratic error term of the j[th] sibling of the i[th] family, $\overline{u}_i$ stands for the average u in family i, and $\overline{u}$ is the grand average of u. Note that the term $*$ is zero as $\sum_{j=1}^{J_i} u_{ij} = \sum_{j=1}^{J_i} \overline{u}_i$. As a result:

$$\sigma_u^2 = \underbrace{\frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{S-1}}_{\substack{\text{within family} \\ \text{variation of } u}} + \underbrace{\frac{\sum_{i=1}^{N} J_i(\overline{u}_i-\overline{u})^2}{S-1}}_{\substack{\text{between family} \\ \text{variation of } u}} \tag{B.9}$$

Now define $\sigma_{u_i}^2$ (the variance of the individual specific error term within family i) :

$$\sigma_{u_i}^2 = \frac{\sum_j (u_{ij}-\overline{u}_i)^2}{J_i - 1} \tag{B.10}$$

Assuming that u_i is homoskedastic and identically distributed in all families, $\sigma_{u_i}^2 = \sigma_u^2 \; \forall i \in \{1, ..., N\}$:

$$\frac{\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{J_i - 1} = \sigma_u^2 \tag{B.11}$$

$$\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2 = (J_i - 1)\sigma_u^2$$

$$\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2 = \sum_{i=1}^{N}(J_i - 1)\sigma_u^2$$

$$\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2 = \sigma_u^2\sum_{i=1}^{N}(J_i - 1)$$

$$\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2 = \sigma_u^2\left[\sum_{i=1}^{N} J_i - N\right]$$

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{S-N}$$

Next, study the terms in parentheses below

$$\begin{aligned}
\epsilon_{ij} &= \alpha_i + u_{ij} \\
\overline{\epsilon}_i &= \alpha_i + \overline{u}_i \\
\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij} - \overline{u}_i)^2 &= \sum_{i=1}^{N}\sum_{j=1}^{J_i}[\epsilon_{ij} - \alpha_i - (\overline{\epsilon}_i - \alpha_i)]^2 \\
\sum_{i=1}^{N}\sum_{j=1}^{J_i}(u_{ij} - \overline{u}_i)^2 &= \sum_{i=1}^{N}\sum_{j=1}^{J_i}(\epsilon_{ij} - \overline{\epsilon}_i)^2
\end{aligned} \tag{B.12}$$

Plugging (B.12) in (B.11) yields:

$$\sigma_u^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i}(\epsilon_{ij} - \overline{\epsilon}_i)^2}{S - N} \tag{B.13}$$

Note that the ANOVA estimator of σ_u^2 in (B.13) uses $S - N$ degrees of freedom, reflecting the pooling of within-family deviations across all N families under the homoskedasticity assumption. This differs from the $S - 1$ denominator in (B.8) and (B.9) , which corresponds to the total sample variance decomposition.

Now, focus on how the deviation of the family specific random term α_i around its sample mean is associated with that of the composite error term and the individual-specific error term below:

$$\begin{aligned}
\overline{\epsilon}_i &= \alpha_i + \overline{u}_i \\
\overline{\epsilon} &= \overline{\alpha} + \overline{u} \\
\alpha_i - \overline{\alpha} &= (\overline{\epsilon}_i - \overline{\epsilon}) - (\overline{u}_i - \overline{u}) \\
\sum_{i=1}^{N} J_i(\alpha_i - \overline{\alpha})^2 &= \sum_{i=1}^{N} J_i[(\overline{\epsilon}_i - \overline{\epsilon}) - (\overline{u}_i - \overline{u})]^2 \\
\sum_{i=1}^{N} J_i(\alpha_i - \overline{\alpha})^2 &= \sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})^2 - 2\underbrace{\sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})(\overline{u}_i - \overline{u})}_{**} + \sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2
\end{aligned} \tag{B.14}$$

Check the term $**$

$$\begin{aligned}
** \equiv &\sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})(\overline{u}_i - \overline{u}) \\
= &\sum_{i=1}^{N} J_i[(\alpha_i - \overline{\alpha}) + (\overline{u}_i - \overline{u})](\overline{u}_i - \overline{u}) \\
= &\underbrace{\sum_{i=1}^{N} J_i(\alpha_i - \overline{\alpha})(\overline{u}_i - \overline{u})}_{\equiv *** = 0} + \sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})(\overline{u}_i - \overline{u}) \\
= &\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2
\end{aligned} \tag{B.15}$$

Since α and u are orthogonal, the term $*** = 0$. As a result,

$$\sum_{i=1}^{N} J_i(\alpha_i - \overline{\alpha})^2 = \sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})^2 - \sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2 \tag{B.16}$$

Adapting (B.13) for σ_α^2 while bearing in mind that since $\alpha_{ij} = \alpha_i$ for $\forall j \in J(i)$, within variation is 0 for σ_α^2 results in

$$\sigma_\alpha^2 = \frac{\sum_{i=1}^{N} J_i(\alpha_i - \overline{\alpha})^2}{S-1} \tag{B.17}$$

Using (B.16) in (B.17) results in

$$\sigma_\alpha^2 = \frac{\sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})^2 - \sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \tag{B.18}$$

where $\overline{\alpha} = \sum_{i=1}^{N} J_i \alpha_i / \sum_{i=1}^{N} J_i$. Now we need to express $\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2$ in terms of σ_u^2.

$$\begin{aligned}
\sigma_u^2 &= \frac{\sum_{i=1}^{N}(J_i - 1)\left(\frac{\sum_{j=1}^{J_i}(u_{ij}-\overline{u}_i)^2}{J_i - 1}\right)}{S-1} + \frac{\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \\
\sigma_u^2 &= \frac{\sum_{i=1}^{N}(J_i - 1)\sigma_u^2}{S-1} + \frac{\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \\
\sigma_u^2 &= \frac{\sigma_u^2(\sum_{i=1}^{N} J_i - N)}{S-1} + \frac{\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \\
\frac{\sigma_u^2(N-1)}{S-1} &= \frac{\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \\
\sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2 &= \sigma_u^2(N-1)
\end{aligned} \tag{B.19}$$

Finally

$$\begin{aligned}
\sigma_\alpha^2 &= \frac{\sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})^2 - \sum_{i=1}^{N} J_i(\overline{u}_i - \overline{u})^2}{S-1} \\
\sigma_\alpha^2 &= \frac{\sum_{i=1}^{N} J_i(\overline{\epsilon}_i - \overline{\epsilon})^2}{S-1} - \frac{\sigma_u^2(N-1)}{S-1}
\end{aligned} \tag{B.20}$$

C. Estimates and Decompositions based on Solon et al. (1991)

Table C.1: Educational sibling correlations and variance components by subgroup: ANOVA estimates

Dimension	Group	σ^2_α Estimate [95% CI]	σ^2_u Estimate [95% CI]	σ^2_ϵ Estimate [95% CI]	ρ Estimate [95% CI]
Whole sample		6.097 [5.833, 6.362]	6.628 [6.297, 6.959]	12.725 [12.402, 13.049]	0.521 [0.501, 0.541]
Birth cohorts					
	1970s	6.931 [6.334, 7.527]	5.906 [5.459, 6.352]	12.836 [12.294, 13.379]	0.540 [0.505, 0.575]
	1970–1980s	7.379 [6.482, 8.276]	7.264 [6.489, 8.039]	14.643 [13.799, 15.488]	0.504 [0.454, 0.553]
	1980s	8.239 [7.572, 8.905]	6.401 [5.944, 6.857]	14.639 [14.026, 15.252]	0.563 [0.531, 0.595]
	1980–1990s	5.268 [4.347, 6.188]	5.272 [4.534, 6.010]	10.540 [9.546, 11.533]	0.500 [0.436, 0.563]
	1990s	4.282 [3.607, 4.958]	5.653 [4.970, 6.335]	9.935 [9.150, 10.720]	0.431 [0.375, 0.487]
Father's education					
	No education	5.073 [4.262, 5.885]	8.084 [7.293, 8.876]	13.158 [12.390, 13.926]	0.386 [0.332, 0.440]
	Primary	5.039 [4.650, 5.428]	6.136 [5.782, 6.489]	11.175 [10.791, 11.559]	0.451 [0.422, 0.480]
	Secondary	3.031 [2.580, 3.481]	4.986 [4.514, 5.458]	8.016 [7.492, 8.541]	0.378 [0.330, 0.426]
	Higher	1.473 [0.957, 1.990]	4.561 [3.800, 5.321]	6.034 [5.235, 6.833]	0.244 [0.166, 0.323]
Mother's education					
	No education	5.585 [5.089, 6.080]	7.347 [6.912, 7.782]	12.932 [12.462, 13.401]	0.432 [0.400, 0.464]
	Primary	4.622 [4.234, 5.010]	5.409 [5.049, 5.770]	10.032 [9.637, 10.426]	0.461 [0.429, 0.493]
	Secondary	3.311 [2.672, 3.949]	3.575 [3.110, 4.041]	6.886 [6.160, 7.612]	0.481 [0.418, 0.543]
	Higher	0.938 [0.022, 1.854]	4.643 [3.083, 6.202]	5.581 [4.074, 7.087]	0.168 [0.006, 0.330]
Age at first fatherhood					
	Below 20s	5.688 [4.256, 7.119]	7.052 [5.973, 8.131]	12.740 [11.280, 14.199]	0.446 [0.366, 0.527]
	20s	6.147 [5.846, 6.448]	5.712 [5.485, 5.940]	11.859 [11.531, 12.187]	0.518 [0.501, 0.536]
	30s	6.983 [6.460, 7.506]	6.577 [6.183, 6.971]	13.559 [13.005, 14.114]	0.515 [0.488, 0.542]
	40s or above	6.876 [5.824, 7.928]	7.157 [6.414, 7.899]	14.033 [13.003, 15.063]	0.490 [0.439, 0.542]
Age at first motherhood					
	Below 20s	5.827 [5.209, 6.444]	5.826 [5.341, 6.312]	11.653 [11.032, 12.274]	0.500 [0.460, 0.540]
	20s	6.487 [6.080, 6.894]	5.966 [5.639, 6.293]	12.453 [12.035, 12.870]	0.521 [0.497, 0.545]
	30s	6.873 [6.003, 7.743]	6.952 [6.199, 7.705]	13.825 [12.956, 14.693]	0.497 [0.447, 0.547]
	40s or above	6.298 [2.534, 10.063]	7.998 [4.766, 11.230]	14.296 [10.986, 17.607]	0.441 [0.219, 0.662]
Sibling gender composition					
	Brothers	6.073 [5.600, 6.546]	5.105 [4.738, 5.473]	11.179 [10.711, 11.646]	0.543 [0.512, 0.574]
	Sisters	10.055 [9.080, 11.030]	5.002 [4.418, 5.586]	15.057 [14.085, 16.029]	0.668 [0.630, 0.706]
	Male/Female	6.058 [5.420, 6.696]	6.786 [6.259, 7.313]	12.844 [12.238, 13.450]	0.472 [0.432, 0.511]
	Female/Male	5.397 [4.728, 6.066]	7.951 [7.308, 8.593]	13.347 [12.646, 14.048]	0.404 [0.362, 0.446]
Sibling count					
	Two siblings	5.649 [5.301, 5.997]	5.207 [4.917, 5.498]	10.857 [10.493, 11.220]	0.520 [0.496, 0.545]
	Three siblings	7.251 [6.549, 7.953]	7.053 [6.464, 7.641]	14.304 [13.653, 14.955]	0.507 [0.468, 0.546]
	Four or more siblings	7.016 [5.745, 8.287]	9.375 [8.327, 10.424]	16.392 [15.234, 17.549]	0.428 [0.365, 0.491]
Sibling age difference					
	0 to 1 year	6.853 [6.166, 7.541]	5.288 [4.786, 5.789]	12.141 [11.450, 12.832]	0.564 [0.525, 0.604]
	2 years	6.415 [5.831, 6.999]	6.295 [5.767, 6.824]	12.710 [12.109, 13.311]	0.505 [0.468, 0.542]
	3 years	6.850 [6.089, 7.611]	5.850 [5.248, 6.451]	12.700 [11.952, 13.447]	0.539 [0.495, 0.584]
	4 years	6.971 [5.998, 7.944]	6.287 [5.535, 7.040]	13.259 [12.322, 14.195]	0.526 [0.471, 0.580]
	5 years or more	6.234 [5.475, 6.994]	6.838 [6.222, 7.454]	13.072 [12.312, 13.832]	0.477 [0.432, 0.521]
Degree of urbanization					
	Urban	6.139 [5.750, 6.528]	5.865 [5.553, 6.177]	12.004 [11.611, 12.398]	0.511 [0.487, 0.536]
	Rural	5.045 [4.519, 5.571]	6.581 [6.083, 7.079]	11.626 [11.092, 12.161]	0.434 [0.396, 0.472]

Notes: Numbers without brackets are point estimates. Numbers in brackets report 95% bootstrap confidence intervals based on 1,500 replications.

Table C.2: Single-layer decomposition of educational sibling correlations: ANOVA estimates

Dimension	Group	γ_k	v_k	w_k	ρ_k	V	W	$\hat{\rho}$	$\tilde{\rho}$	ρ	π
Birth cohorts	1970s	0.2688	1.8632	0.0002	0.5438	6.6275	0.0008	0.5209	0.3397	0.5209	-0.0010
	1970–1980s	0.1277	0.9421	0.0001	0.5047						
	1980s	0.2852	2.3495	0.0002	0.5675						
	1980–1990s	0.1123	0.5917	0.0001	0.5110						
	1990s	0.2057	0.8810	0.0002	0.4340						
Father's education	No education	0.1653	0.8386	0.3750	0.3856	4.3606	2.2678	0.4170	0.9995	0.5209	0.2490
	Primary	0.5450	2.7462	1.2359	0.4509						
	Secondary	0.2242	0.6795	0.5086	0.3780						
	Higher	0.0653	0.0963	0.1483	0.2442						
Mother's education	No education	0.4408	2.4616	0.7741	0.4319	4.8723	1.7561	0.4442	0.9995	0.5209	0.1726
	Primary	0.4583	2.1182	0.8048	0.4608						
	Secondary	0.0834	0.2762	0.1466	0.4808						
	Higher	0.0173	0.0162	0.0305	0.1681						
First fatherhood age	Below 20s	0.0345	0.2077	0.0093	0.4677	6.3591	0.2693	0.5106	0.9953	0.5209	0.0202
	20s	0.6345	3.8617	0.1709	0.5140						
	30s	0.2676	1.8536	0.0721	0.5126						
	40s or above	0.0632	0.4362	0.0170	0.4944						
First motherhood age	Below 20s	0.2127	1.2393	0.0490	0.5000	6.3978	0.2305	0.5121	0.9944	0.5209	0.0172
	20s	0.6352	4.1204	0.1464	0.5209						
	30s	0.1412	0.9707	0.0326	0.4972						
	40s or above	0.0107	0.0674	0.0025	0.4406						
Sibling gender composition	Brothers	0.3674	2.2315	0.0075	0.5433	6.6079	0.0204	0.5201	0.9450	0.5209	0.0014
	Sisters	0.1692	1.7017	0.0035	0.6678						
	Male/Female	0.2650	1.6056	0.0054	0.4717						
	Female/Male	0.1981	1.0692	0.0040	0.4043						
Sibling count	Two	0.6515	3.6805	0.2912	0.5203	6.1814	0.4469	0.5035	0.9978	0.5209	0.0346
	Three	0.2422	1.7564	0.1083	0.5069						
	Four or more	0.1061	0.7446	0.0475	0.4280						
Sibling age difference	0 to 1 year	0.2137	1.4646	0.0017	0.5645	6.6202	0.0081	0.5206	0.8386	0.5209	0.0005
	2 years	0.2981	1.9123	0.0024	0.5047						
	3 years	0.1838	1.2591	0.0015	0.5394						
	4 years	0.1197	0.8344	0.0010	0.5258						
	5 years or more	0.1844	1.1498	0.0015	0.4769						
Degree of urbanization	Urban	0.6761	4.1506	0.5706	0.5114	5.7844	0.8440	0.4869	0.9995	0.5209	0.0699
	Rural	0.3238	1.6337	0.2733	0.4339						

Notes: The table reports the single-layer decomposition results based on the ANOVA estimates obtained following Solon et al. (1991). V, W, $\hat{\rho}$, $\tilde{\rho}$, ρ, and π are partition-level quantities and are therefore shown once for each partition.

Table C.3: Subset-removal counterfactual decompositions for paternal and maternal education (REML estimates)

Excluded subset	Remaining subset	$\gamma_{k,-m}$	v_k	w_k	V_{-m}	W_{-m}	ρ_k	$\tilde{\rho}_{-m}$	$\hat{\rho}_{-m}$	ρ_{-m}	π_{-m}	ρ	$\hat{\rho}$	$\tilde{\rho}$	V	W	π	ρ^c_{-m}	π^c_{-m}	$\dot{\Delta}_m$	$\ddot{\Delta}_m$	Δ_m
Panel A: Father's education																						
	Primary	0.653	3.292	0.971			0.452															
No education	Secondary	0.269	0.772	0.400	4.140	1.487	0.377	0.930	0.428	0.499	0.166							0.501	0.171	-0.023	-0.002	-0.026
	Higher	0.078	0.077	0.116			0.226															
	No education	0.363	1.758	1.860			0.373															
Primary	Secondary	0.493	1.416	2.523	3.314	5.118	0.377	0.971	0.365	0.587	0.610							0.584	0.603	0.060	0.003	0.062
	Higher	0.144	0.141	0.735			0.226					0.525	0.416	0.959	4.255	2.443	0.260					
	No education	0.213	1.031	0.445			0.373															
Secondary	Primary	0.703	3.542	1.465	4.655	2.086	0.452	0.952	0.424	0.512	0.207							0.513	0.208	-0.012	-0.001	-0.012
	Higher	0.084	0.083	0.176			0.226															
	No education	0.177	0.856	0.311			0.373															
Higher	Primary	0.583	2.940	1.025	4.484	1.757	0.452	0.975	0.422	0.502	0.190							0.501	0.187	-0.024	0.001	-0.023
	Secondary	0.240	0.689	0.422			0.377															
Panel B: Mother's education																						
	Primary	0.820	3.799	0.603			0.464															
No education	Secondary	0.149	0.390	0.110	4.202	0.735	0.473	0.900	0.460	0.497	0.078							0.499	0.084	-0.026	-0.002	-0.028
	Higher	0.031	0.012	0.023			0.116															
	No education	0.814	4.508	2.235			0.430															
Primary	Secondary	0.154	0.403	0.423	4.923	2.747	0.473	0.952	0.431	0.536	0.244							0.536	0.246	0.012	-0.001	0.011
	Higher	0.032	0.013	0.088			0.116					0.525	0.445	0.958	4.791	1.907	0.180					
	No education	0.481	2.664	0.641			0.430															
Secondary	Primary	0.500	2.318	0.666	4.990	1.333	0.464	0.965	0.443	0.500	0.129							0.500	0.128	-0.025	0.000	-0.024
	Higher	0.019	0.008	0.025			0.116															
	No education	0.449	2.485	0.764			0.430															
Higher	Primary	0.466	2.162	0.795	4.869	1.704	0.464	0.965	0.446	0.519	0.162							0.518	0.161	-0.006	0.000	-0.006
	Secondary	0.085	0.222	0.145			0.473															

Notes: The table reports subset-removal counterfactual decompositions for the paternal- and maternal-education partitioning schemes based on REML estimates.

Table C.4: Subset-removal counterfactual decompositions for paternal and maternal education (ANOVA estimates)

Excluded subset	*Remaining subset	$\gamma_{k,-m}$	v_k	w_k	V_{-m}	W_{-m}	ρ_k	$\tilde{\rho}_{-m}$	$\hat{\rho}_{-m}$	ρ_{-m}	π_{-m}	ρ	$\hat{\rho}$	$\tilde{\rho}$	V	W	π	ρ^c_{-m}	π^c_{-m}	$\dot{\Delta}_m$	$\ddot{\Delta}_m$	Δ_m
Panel A: Father's education																						
	Primary	0.6529	3.2903	0.8867			0.451															
No education	Secondary	0.2687	0.8142	0.3649	4.220	1.358	0.378	0.999	0.425	0.494	0.163							0.494	0.163	-0.026	0.000	-0.026
	Higher	0.0783	0.1153	0.1064			0.244															
	No education	0.3633	1.8432	1.7130			0.386															
Primary	Secondary	0.4928	1.4935	2.3233	3.548	4.714	0.378	1.000	0.370	0.577	0.561							0.577	0.561	0.056	0.000	0.056
	Higher	0.1436	0.2115	0.6773			0.244					0.521	0.417	1.000	4.361	2.268	0.249					
	No education	0.2131	1.0811	0.4210			0.386															
Secondary	Primary	0.7025	3.5402	1.3874	4.745	1.975	0.451	0.999	0.425	0.511	0.203							0.511	0.203	-0.009	0.000	-0.009
	Higher	0.0842	0.1241	0.1665			0.244															
	No education	0.1769	0.8973	0.2887			0.386															
Higher	Primary	0.5831	2.9383	0.9516	4.563	1.632	0.451	0.999	0.424	0.500	0.179							0.500	0.179	-0.021	0.000	-0.021
	Secondary	0.2399	0.7271	0.3916			0.378															
Panel B: Mother's education																						
	Primary	0.8196	3.7883	0.5192			0.461															
No education	Secondary	0.1492	0.4940	0.0946	4.311	0.633	0.481	0.998	0.458	0.492	0.075							0.492	0.075	-0.029	0.000	-0.029
	Higher	0.0309	0.0290	0.0197			0.168															
	No education	0.8138	4.5446	2.0435			0.432															
Primary	Secondary	0.1541	0.5100	0.3871	5.085	2.511	0.481	1.000	0.432	0.532	0.231							0.532	0.231	0.011	0.000	0.011
	Higher	0.0319	0.0300	0.0805			0.168					0.521	0.444	0.999	4.872	1.756	0.173					
	No education	0.4810	2.6859	0.6091			0.432															
Secondary	Primary	0.5000	2.3112	0.6333	5.015	1.266	0.461	0.999	0.442	0.498	0.127							0.498	0.127	-0.023	0.000	-0.023
	Higher	0.0189	0.0177	0.0240			0.168															
	No education	0.4486	2.5051	0.7033			0.432															
Higher	Primary	0.4664	2.1556	0.7312	4.942	1.568	0.461	1.000	0.447	0.515	0.154							0.515	0.154	-0.006	0.000	-0.006
	Secondary	0.0849	0.2811	0.1332			0.481															

Notes: The table reports subset-removal counterfactual decompositions for the paternal- and maternal-education partitioning schemes based on the ANOVA estimator. The final verification column from the spreadsheet is omitted. In these ANOVA-based counterfactuals, the between-channel contribution $\ddot{\Delta}_m$ is effectively zero throughout, so the subset-removal effects operate almost entirely through the within-channel.

Table C.5: Full dual-layer decomposition results by second-layer partition (REML estimates)

Second-layer partition	Subsubset	Gender profile	ϕ	v	w	$\rho_{k,l}$	$\tilde{\rho}_k$	V_k	$\hat{\rho}_k$	W_k	ρ_k	π_k
Birth cohorts												
	1970s	Brothers	0.273	1.904	-0.007	0.589	-3.199	6.064	0.546	-0.026	0.543	-0.005
	1970–1980s	Brothers	0.133	0.810	-0.003	0.492	-3.199	6.064	0.546	-0.026	0.543	-0.005
	1980s	Brothers	0.283	2.019	-0.007	0.618	-3.199	6.064	0.546	-0.026	0.543	-0.005
	1980–1990s	Brothers	0.127	0.663	-0.003	0.566	-3.199	6.064	0.546	-0.026	0.543	-0.005
	1990s	Brothers	0.182	0.667	-0.005	0.373	-3.199	6.064	0.546	-0.026	0.543	-0.005
	1970s	Sisters	0.263	2.625	0.004	0.677	0.100	10.051	0.677	0.015	0.671	-0.009
	1970–1980s	Sisters	0.103	1.243	0.002	0.702	0.100	10.051	0.677	0.015	0.671	-0.009
	1980s	Sisters	0.295	3.993	0.004	0.728	0.100	10.051	0.677	0.015	0.671	-0.009
	1980–1990s	Sisters	0.092	0.610	0.001	0.571	0.100	10.051	0.677	0.015	0.671	-0.009
	1990s	Sisters	0.244	1.580	0.004	0.596	0.100	10.051	0.677	0.015	0.671	-0.009
	1970s	Male/Female	0.269	1.751	0.020	0.508	0.345	6.517	0.517	0.075	0.514	-0.006
	1970–1980s	Male/Female	0.142	1.204	0.011	0.538	0.345	6.517	0.517	0.075	0.514	-0.006
	1980s	Male/Female	0.264	2.222	0.020	0.562	0.345	6.517	0.517	0.075	0.514	-0.006
	1980–1990s	Male/Female	0.124	0.666	0.009	0.529	0.345	6.517	0.517	0.075	0.514	-0.006
	1990s	Male/Female	0.200	0.673	0.015	0.395	0.345	6.517	0.517	0.075	0.514	-0.006
	1970s	Female/Male	0.264	1.774	-0.006	0.507	-0.879	5.864	0.439	-0.022	0.436	-0.006
	1970–1980s	Female/Male	0.118	0.883	-0.003	0.515	-0.879	5.864	0.439	-0.022	0.436	-0.006
	1980s	Female/Male	0.307	2.104	-0.007	0.431	-0.879	5.864	0.439	-0.022	0.436	-0.006
	1980–1990s	Female/Male	0.085	0.349	-0.002	0.324	-0.879	5.864	0.439	-0.022	0.436	-0.006
	1990s	Female/Male	0.224	0.754	-0.005	0.343	-0.879	5.864	0.439	-0.022	0.436	-0.006
Father's education												
	No education	Brothers	0.181	1.111	0.311	0.514	0.923	4.326	0.467	1.712	0.543	0.163
	Primary	Brothers	0.562	2.563	0.963	0.476	0.923	4.326	0.467	1.712	0.543	0.163
	Secondary	Brothers	0.210	0.593	0.359	0.382	0.923	4.326	0.467	1.712	0.543	0.163
	Higher	Brothers	0.046	0.058	0.080	0.359	0.923	4.326	0.467	1.712	0.543	0.163
	No education	Sisters	0.139	1.033	0.484	0.632	0.911	6.591	0.589	3.477	0.671	0.139
	Primary	Sisters	0.509	4.247	1.770	0.617	0.911	6.591	0.589	3.477	0.671	0.139
	Secondary	Sisters	0.256	1.114	0.890	0.555	0.911	6.591	0.589	3.477	0.671	0.139
	Higher	Sisters	0.095	0.196	0.333	0.296	0.911	6.591	0.589	3.477	0.671	0.139
	No education	Male/Female	0.166	0.701	0.443	0.327	0.978	3.921	0.389	2.672	0.514	0.323
	Primary	Male/Female	0.550	2.608	1.469	0.426	0.978	3.921	0.389	2.672	0.514	0.323
	Secondary	Male/Female	0.221	0.573	0.591	0.348	0.978	3.921	0.389	2.672	0.514	0.323
	Higher	Male/Female	0.063	0.039	0.169	0.218	0.978	3.921	0.389	2.672	0.514	0.323
	No education	Female/Male	0.157	0.371	0.404	0.174	0.937	3.275	0.307	2.568	0.436	0.419
	Primary	Female/Male	0.537	2.472	1.379	0.376	0.937	3.275	0.307	2.568	0.436	0.419

Continued on next page

Table C.5: Full dual-layer decomposition results by second-layer partition (REML estimates) — continued

Second-layer partition	Subsubset	Gender profile	ϕ	v	w	$\rho_{k,l}$	$\tilde{\rho}_k$	V_k	$\hat{\rho}_k$	W_k	ρ_k	π_k
	Secondary	Female/Male	0.228	0.416	0.586	0.248	0.937	3.275	0.307	2.568	0.436	0.419
	Higher	Female/Male	0.077	0.017	0.199	0.060	0.937	3.275	0.307	2.568	0.436	0.419
Mother's education												
	No education	Brothers	0.475	2.651	0.561	0.517	0.885	4.856	0.496	1.182	0.543	0.094
	Primary	Brothers	0.431	1.954	0.510	0.469	0.885	4.856	0.496	1.182	0.543	0.094
	Secondary	Brothers	0.080	0.241	0.095	0.522	0.885	4.856	0.496	1.182	0.543	0.094
	Higher	Brothers	0.013	0.010	0.016	0.344	0.885	4.856	0.496	1.182	0.543	0.094
	No education	Sisters	0.384	3.591	1.208	0.625	0.925	6.921	0.597	3.147	0.671	0.125
	Primary	Sisters	0.485	3.073	1.526	0.592	0.925	6.921	0.597	3.147	0.671	0.125
	Secondary	Sisters	0.103	0.244	0.325	0.462	0.925	6.921	0.597	3.147	0.671	0.125
	Higher	Sisters	0.027	0.013	0.087	0.093	0.925	6.921	0.597	3.147	0.671	0.125
	No education	Male/Female	0.443	2.134	0.909	0.361	0.990	4.540	0.422	2.052	0.514	0.217
	Primary	Male/Female	0.471	2.249	0.966	0.499	0.990	4.540	0.422	2.052	0.514	0.217
	Secondary	Male/Female	0.070	0.157	0.143	0.495	0.990	4.540	0.422	2.052	0.514	0.217
	Higher	Male/Female	0.016	0.000	0.033	0.000	0.990	4.540	0.422	2.052	0.514	0.217
	No education	Female/Male	0.423	2.197	0.849	0.373	0.856	3.837	0.347	2.006	0.436	0.256
	Primary	Female/Male	0.469	1.477	0.940	0.315	0.856	3.837	0.347	2.006	0.436	0.256
	Secondary	Female/Male	0.091	0.157	0.182	0.355	0.856	3.837	0.347	2.006	0.436	0.256
	Higher	Female/Male	0.017	0.006	0.035	0.244	0.856	3.837	0.347	2.006	0.436	0.256
First fatherhood age												
	below_20s	Brothers	0.037	0.162	0.009	0.483	1.101	5.792	0.532	0.246	0.543	0.021
	20s	Brothers	0.637	3.672	0.157	0.534	1.101	5.792	0.532	0.246	0.543	0.021
	30s	Brothers	0.264	1.587	0.065	0.535	1.101	5.792	0.532	0.246	0.543	0.021
	40s or above	Brothers	0.062	0.371	0.015	0.523	1.101	5.792	0.532	0.246	0.543	0.021
	below_20s	Sisters	0.034	0.238	0.024	0.459	1.016	9.372	0.655	0.694	0.671	0.025
	20s	Sisters	0.635	5.238	0.441	0.634	1.016	9.372	0.655	0.694	0.671	0.025
	30s	Sisters	0.262	3.117	0.182	0.705	1.016	9.372	0.655	0.694	0.671	0.025
	40s or above	Sisters	0.068	0.779	0.047	0.698	1.016	9.372	0.655	0.694	0.671	0.025
	below_20s	Male/Female	0.037	0.222	0.010	0.372	1.524	6.331	0.500	0.261	0.514	0.027
	20s	Male/Female	0.631	3.858	0.164	0.520	1.524	6.331	0.500	0.261	0.514	0.027
	30s	Male/Female	0.272	1.902	0.071	0.494	1.524	6.331	0.500	0.261	0.514	0.027
	40s or above	Male/Female	0.060	0.348	0.016	0.444	1.524	6.331	0.500	0.261	0.514	0.027
	below_20s	Female/Male	0.027	0.199	0.007	0.493	1.037	5.585	0.425	0.257	0.436	0.027
	20s	Female/Male	0.635	3.372	0.163	0.422	1.037	5.585	0.425	0.257	0.436	0.027
	30s	Female/Male	0.272	1.479	0.070	0.407	1.037	5.585	0.425	0.257	0.436	0.027
	40s or above	Female/Male	0.066	0.535	0.017	0.480	1.037	5.585	0.425	0.257	0.436	0.027

Continued on next page

Table C.5: Full dual-layer decomposition results by second-layer partition (REML estimates) — continued

Second-layer partition	Subsubset	Gender profile	ϕ	v	w	$\rho_{k,l}$	$\tilde{\rho}_k$	V_k	$\hat{\rho}_k$	W_k	ρ_k	π_k
First motherhood age												
	below_20s	Brothers	0.224	1.324	0.026	0.579	0.869	5.920	0.539	0.118	0.543	0.007
	20s	Brothers	0.634	3.652	0.075	0.518	0.869	5.920	0.539	0.118	0.543	0.007
	30s	Brothers	0.130	0.911	0.015	0.584	0.869	5.920	0.539	0.118	0.543	0.007
	40s or above	Brothers	0.012	0.033	0.001	0.409	0.869	5.920	0.539	0.118	0.543	0.007
	below_20s	Sisters	0.197	1.647	0.120	0.658	0.895	9.457	0.660	0.609	0.671	0.016
	20s	Sisters	0.644	6.216	0.393	0.662	0.895	9.457	0.660	0.609	0.671	0.016
	30s	Sisters	0.151	1.588	0.092	0.668	0.895	9.457	0.660	0.609	0.671	0.016
	40s or above	Sisters	0.007	0.007	0.005	0.141	0.895	9.457	0.660	0.609	0.671	0.016
	below_20s	Male/Female	0.222	1.295	0.054	0.461	1.060	6.348	0.504	0.244	0.514	0.020
	20s	Male/Female	0.625	3.997	0.153	0.526	1.060	6.348	0.504	0.244	0.514	0.020
	30s	Male/Female	0.143	0.990	0.035	0.488	1.060	6.348	0.504	0.244	0.514	0.020
	40s or above	Male/Female	0.009	0.066	0.002	0.433	1.060	6.348	0.504	0.244	0.514	0.020
	below_20s	Female/Male	0.193	0.832	0.101	0.356	1.110	5.322	0.412	0.520	0.436	0.059
	20s	Female/Male	0.642	3.667	0.334	0.446	1.110	5.322	0.412	0.520	0.436	0.059
	30s	Female/Male	0.151	0.796	0.079	0.373	1.110	5.322	0.412	0.520	0.436	0.059
	40s or above	Female/Male	0.012	0.027	0.007	0.120	1.110	5.322	0.412	0.520	0.436	0.059
Sibling count												
	Two siblings	Brothers	0.622	3.298	0.202	0.523	0.911	5.713	0.531	0.325	0.543	0.023
	Three siblings	Brothers	0.260	1.601	0.084	0.542	0.911	5.713	0.531	0.325	0.543	0.023
	Four or more siblings	Brothers	0.118	0.813	0.038	0.540	0.911	5.713	0.531	0.325	0.543	0.023
	Two siblings	Sisters	0.645	4.987	0.329	0.641	0.812	9.557	0.665	0.509	0.671	0.009
	Three siblings	Sisters	0.244	3.221	0.124	0.741	0.812	9.557	0.665	0.509	0.671	0.009
	Four or more siblings	Sisters	0.110	1.349	0.056	0.599	0.812	9.557	0.665	0.509	0.671	0.009
	Two siblings	Male/Female	0.676	3.827	0.451	0.535	0.924	5.925	0.490	0.667	0.514	0.050
	Three siblings	Male/Female	0.231	1.481	0.154	0.451	0.924	5.925	0.490	0.667	0.514	0.050
	Four or more siblings	Male/Female	0.093	0.618	0.062	0.370	0.924	5.925	0.490	0.667	0.514	0.050
	Two siblings	Female/Male	0.678	3.103	0.478	0.440	0.854	5.137	0.409	0.705	0.436	0.067
	Three siblings	Female/Male	0.224	1.526	0.158	0.400	0.854	5.137	0.409	0.705	0.436	0.067
	Four or more siblings	Female/Male	0.098	0.508	0.069	0.298	0.854	5.137	0.409	0.705	0.436	0.067
Sibling age difference												
	0 to 1 year	Brothers	0.196	1.044	0.000	0.548	-0.204	6.039	0.544	-0.001	0.543	-0.001
	2 years	Brothers	0.302	1.945	0.000	0.569	-0.204	6.039	0.544	-0.001	0.543	-0.001
	3 years	Brothers	0.186	1.182	0.000	0.589	-0.204	6.039	0.544	-0.001	0.543	-0.001
	4 years	Brothers	0.121	0.792	0.000	0.558	-0.204	6.039	0.544	-0.001	0.543	-0.001

Continued on next page

Table C.5: Full dual-layer decomposition results by second-layer partition (REML estimates) — continued

Second-layer partition	**Subsubset**	**Gender profile**	ϕ	v	w	$\rho_{k,l}$	$\tilde{\rho}_k$	V_k	$\hat{\rho}_k$	W_k	ρ_k	π_k
	5 years or more	Brothers	0.194	1.077	0.000	0.455	-0.204	6.039	0.544	-0.001	0.543	-0.001
	0 to 1 year	Sisters	0.249	2.599	0.028	0.736	0.586	9.955	0.672	0.112	0.671	-0.002
	2 years	Sisters	0.305	3.089	0.034	0.654	0.586	9.955	0.672	0.112	0.671	-0.002
	3 years	Sisters	0.185	1.946	0.021	0.741	0.586	9.955	0.672	0.112	0.671	-0.002
	4 years	Sisters	0.107	1.109	0.012	0.655	0.586	9.955	0.672	0.112	0.671	-0.002
	5 years or more	Sisters	0.152	1.213	0.017	0.541	0.586	9.955	0.672	0.112	0.671	-0.002
	0 to 1 year	Male/Female	0.165	1.124	-0.024	0.556	2.422	6.739	0.523	-0.147	0.514	-0.017
	2 years	Male/Female	0.270	1.666	-0.040	0.496	2.422	6.739	0.523	-0.147	0.514	-0.017
	3 years	Male/Female	0.191	1.154	-0.028	0.455	2.422	6.739	0.523	-0.147	0.514	-0.017
	4 years	Male/Female	0.147	1.205	-0.022	0.577	2.422	6.739	0.523	-0.147	0.514	-0.017
	5 years or more	Male/Female	0.227	1.591	-0.033	0.552	2.422	6.739	0.523	-0.147	0.514	-0.017
	0 to 1 year	Female/Male	0.281	1.809	0.014	0.470	0.396	5.794	0.437	0.048	0.436	-0.001
	2 years	Female/Male	0.323	1.577	0.016	0.372	0.396	5.794	0.437	0.048	0.436	-0.001
	3 years	Female/Male	0.169	1.211	0.008	0.513	0.396	5.794	0.437	0.048	0.436	-0.001
	4 years	Female/Male	0.090	0.473	0.004	0.429	0.396	5.794	0.437	0.048	0.436	-0.001
	5 years or more	Female/Male	0.137	0.723	0.007	0.421	0.396	5.794	0.437	0.048	0.436	-0.001
Degree of urbanization												
	Urban	Brothers	0.656	3.932	0.372	0.551	0.897	5.471	0.522	0.567	0.543	0.041
	Rural	Brothers	0.344	1.539	0.195	0.459	0.897	5.471	0.522	0.567	0.543	0.041
	Urban	Sisters	0.696	5.949	1.094	0.626	0.973	8.495	0.635	1.572	0.671	0.057
	Rural	Sisters	0.304	2.546	0.478	0.655	0.973	8.495	0.635	1.572	0.671	0.057
	Urban	Male/Female	0.668	4.072	0.564	0.506	0.983	5.748	0.480	0.844	0.514	0.070
	Rural	Male/Female	0.332	1.676	0.280	0.427	0.983	5.748	0.480	0.844	0.514	0.070
	Urban	Female/Male	0.706	3.700	0.670	0.434	0.905	4.894	0.397	0.948	0.436	0.100
	Rural	Female/Male	0.293	1.194	0.278	0.313	0.905	4.894	0.397	0.948	0.436	0.100

Notes: The table reports the full REML-based dual-layer decomposition results underlying the compact main-text summaries. The second-layer partition identifies the observed shared family characteristic used to split each top-level gender-based subset. For readability, the subset-level objects $\tilde{\rho}_k$, V_k, $\hat{\rho}_k$, W_k, ρ_k, and π_k are repeated within each gender-profile block. Corresponding ANOVA-based results are extremely similar and are available upon request.