The Total Effect of Social Origins on Educational Attainment: Meta-Analysis of Sibling Correlations from 18 Countries*

Lewis R. Anderson

Patrick Präg

Evelina T. Akimova

Christiaan Monden

July 10, 2024

Abstract

The sibling correlation (SC) estimates the total effect of family background, or 'social origins,' and may be interpreted as measuring a society's inequality of opportunity. Its sensitivity to both observed and unobserved factors makes it an all-encompassing measure and an attractive choice for comparative research. We gather and summarise all available estimates of SCs in educational attainment (M = .46, SD = .09), before employing meta-regression to explore variability in these estimates. First, we find significantly lower SCs in Sweden, Norway, Finland, and Denmark than the US, with US correlations on the order of 0.1—or 25%—higher. Most of the other (primarily European) countries for which we find estimates fall in between. Second, we find a novel Great Gatsby Curve-type positive association between income inequality in childhood and the SC, both cross-nationally and within countries over time. This supports theoretical accounts of the Great Gatsby Curve that emphasize the role of educational inequality as a link between economic inequality and social

*This manuscript is forthcoming at *Demography*. Author affiliations on the last page.
Previous versions of this study were presented at the 2019 Annual Meeting of the Population Association of America in Austin, the 2019 RC28 Spring Meeting in Frankfurt, the 2024 GEOM conference in Bari, and a workshop at Nuffield College in 2019 as well as seminars at Trinity College Dublin, University College London, University of Oslo, and University of Oxford. We thank discussants at these meetings as well as Andreas Haupt, Bastian Betthäuser, Betsy Becker, Brian Nolan, Elizabeth Tipton, John Ermisch, John Goldthorpe, Kristian Bernt

immobility, and implies that greater equality of educational opportunity likely requires a reduction in economic inequality. Additionally, we find that correlations between sisters are modestly higher on average, and we find no overall differences between cohorts.

Introduction

Equality of opportunity is a widely-held ideal and an explicit goal of much policymaking. Comparing societies to better understand the factors that facilitate or hinder this has long been a major research area for social scientists. To measure equality of opportunity, researchers mostly estimate some form of association between parent and child socio-economic position in terms of a single indicator (Breen and Müller, 2020; Chetty *et al.*, 2017; Hout, 2018; Sorokin, 1927). However, operationalizing family background in this way underestimates its effects. Not only do multiple dimensions of parental socio-economic position play a role in the intergenerational transmission of advantage (Bukodi and Goldthorpe, 2013), but so too do a plethora of other difficult-to-observe parental characteristics (Lareau, 2011; Putnam, 2015).

This study focuses on an alternative measure, namely the sibling correlation (SC)—the degree of similarity between siblings on a given outcome. The SC corresponds to the proportion of variance attributable to factors that increase sibling resemblance. The major dimensions of parental socio-economic position—income, wealth, occupation, and education—all act in this way and their effects are thus captured without needing to be measured, as too are the effects of other family-level factors such as ethnicity, and important unobservables such as parental motivation. Accordingly, the SC estimates the total effect of social origins. This makes it a valuable measure of a society's inequality of opportunity, albeit the theoretical correspondence is more nuanced than is typically appreciated. We provide a reinterpretation of the link between the SC and inequality of opportunity, and compare this measure's strengths and limitations to alternatives.

We focus on the SC in *education*, a highly important pathway in the transmission of advantage from one generation to the next and a key predictor of children's life chances. In our study, we compare SCs in education across countries and birth cohorts. We report the results of a systematic literature search for all published SCs in educational attainment. Next, we employ meta-regression

Karlson, Noel Card, Per Engzell, Richard Breen, Thees Spreckelsen, Tobias Rüttenauer, and the four anonymous reviewers at *Demography*.

Replication materials for this study are available on-line (Anderson *et al.*, 2024).

This study has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement no. 681546 (FAMSIZEMATTERS) and grant agreement no. 856455 (DINA). Patrick Präg's work is supported by a grant of the French National Research Agency ANR, 'Investissements

to describe cross-national and temporal variation, and then to test whether equality of opportunity systematically differs across contexts characterised by different levels of income inequality.

Bringing together SCs from the broadest possible range of country–cohort contexts allows us to contribute new insights to current debates on cross-national differences in inequality of educational opportunity (Grätz *et al.*, 2021; Hertz *et al.*, 2008) and the idea that more economically unequal societies exhibit greater intergenerational persistence (Blanden, 2013; Corak, 2013; Jerrim and Macmillan, 2015)—the so-called 'Great Gatsby Curve.' Prior research using SCs in education mostly comprises single-country studies. A few studies have compared SCs in education across countries (Grätz *et al.*, 2021; Sieben *et al.*, 2001; Sieben and de Graaf, 2001, 2003; Sirniö *et al.*, 2020), but these investigations have not yet succeeded in 'replacing the notion that "nations differ" by statements formulated in terms of specific variables' (Przeworski and Teune, 1970, p. 29–30). By investigating the role of income inequality, we build upon this prior work to enhance our understanding of the correlates of a society's opportunity structure.

We proceed by first discussing SCs in detail. Next, we elaborate on the relevance of income inequality. Then we present our empirical analysis.

Sibling correlations

This section aims to clarify what SCs measure, in what sense this corresponds to inequality of opportunity, and how this complements two alternative measures of (in)equality of opportunity, genetically-sensitive variance decompositions, and parent–child associations.

What sibling correlations measure

Most straightforwardly, SCs measure how strongly siblings in a population resemble one another on an outcome. A value of *zero* means that siblings are as similar as randomly-chosen individuals of the same birth cohort, in expectation; *one* means that siblings in a family all share the same outcome—from which it is easy to intuit that family background likely plays an important role.

SCs are usually estimated in one of two ways. The first is simply the Pearson correlation between the outcomes for two randomly-selected siblings within each family. For intuition, imagine plotting sibling A's outcome on one axis and sibling B's on the other for a sample of families. An extension yielding the same

d'Avenir' (LabEx Ecodec/ANR-11-LABX-0047). Evelina Akimova's work is supported by a grant of the Leverhulme Trust (no. RC-2018-003) for the Leverhulme Center for Demographic Science and the European Research Council (ERC) no. 835079 (CHRONO).

expected value is to include all sibling pairs and weight each family equally to account for differences in sibship size.

Second, SCs may be estimated within a variance-components framework. The outcome Y for individual i is decomposed into an individual-specific component S_{if} and a family component C_f common to siblings in the same family f:

$$Y_{if} = S_{if} + C_f \tag{1}$$

S_{if} and C_f are thus assumed uncorrelated (Conley and Glauber, 2005; Lundberg, 2020). Variance in the outcome is decomposed into a between-family component σ_C^2 and a within-family component σ_S^2. The SC is given by the intraclass correlation coefficient $\frac{\sigma_C^2}{\sigma_S^2+\sigma_C^2}$, i.e. the proportion of the variance that is *between* rather than *within* families.

SCs, then, measure the proportion of the variance in an outcome attributable to factors whose effect is to increase similarity among siblings (on that outcome).[1] Insofar as they increase sibling similarity, then, influences captured by the SC include all relevant parental characteristics and behaviours, and a family's economic, social, and cultural resources. They further include school, community, and neighbourhood factors as well as extra-nuclear kin. The SC also captures the effect of the, on average, 50 % of genes that siblings share. Together we term these factors a sibling set's 'common inheritance,' to emphasize both that environmental and genetic factors are included, and that the remaining variance is explained by individual-specific factors, albeit these may be rooted in the family. For consistency with the wider literature, we use the terms 'social origins' and 'family background' interchangeably with 'common inheritance', albeit we argue the latter provides a more precise descriptor.

Correspondingly, *one minus the sibling correlation* $(1 - SC)$ gives the variance explained by factors that generate differences between siblings and thereby bring their similarity closer to that between random individuals. These 'individualizing factors,' as we term them, arise from the corresponding non-shared 50 % of genes and from non-shared environmental effects (Plomin, 2011). Sources of the latter include birth order and spacing (Barclay and Smith, 2022; Black *et al.*, 2005; Chan *et al.*, 2019), maternal age (Grätz and Wiborg, 2024), sex (Breen *et al.*, 2010; Buchmann *et al.*, 2008), and non-systematic sibling-specific experiences such as random allocation to a particular schoolteacher. Importantly, sibling differences are also rooted in differential experience of what might normally be understood as the shared environment (Conley, 2008; Grätz, 2018). Families, schools, and communities may change over time, varying the age-specific exposures of different siblings. As discussed further below, siblings may be treated

[1] Note that this differs from the formulation 'factors shared by siblings.' Siblings may be affected differently by factors they share (Maccoby, 2000), as we discuss below.

differently by parents and teachers, and fall into different peer groups.[2]

Siblings are not just the passive recipients of differences in their individual-specific environments, but also generate these differences in a process of dynamic interplay. Parents respond to initial differences in early temperament with differential parenting (Bates *et al.*, 2012; Hernández-Alava and Popli, 2017; Huh *et al.*, 2006; Kiff *et al.*, 2011). Children's genetic propensity to educational attainment elicits more of some forms of cognitively-stimulating parenting in early childhood (Breinholt and Conley, 2023). Siblings actively seek to differentiate themselves from one another and occupy a unique niche, consistent with evolutionary theory's prediction that this strategy increases parental investment (Healey and Ellis, 2007; Sulloway, 1996). In this light, some sources of within-family inequality can be viewed not as inexorable forces but as circumstances to which children respond as individual agents, shaping their own micro-environment.

Sibling correlations and inequality of opportunity

How does the foregoing relate to inequality of opportunity? We first critique the conventional answer before providing our own.

Both answers share the point of departure that the SC captures the effect of social origins, and one's social origins are beyond one's control, meaning (following, among others, Roemer, 1998) that one cannot be held accountable for their effects, whether positive or negative. Inequality of opportunity in this context refers to the extent to which individuals' outcomes are determined by factors beyond their control. However, invoking equality of opportunity requires an account of what constitutes merit, the set of characteristics for which one can be held accountable, and which determine success under a regime of equality of opportunity.

The SC is conventionally characterized as a *lower-bound estimate* of the total effect of social origins (Björklund and Jäntti, 2012; Schnitzlein, 2014). 'Lower-bound' because factors such as birth order and sibling-specific parenting that increase sibling differences are explicitly framed *as family background effects* which are however missed (Bredtmann and Smith, 2018; Halliday and Mazumder, 2017; Karhula *et al.*, 2019; Lillehagen and Isungset, 2020).

As a link between the SC and inequality of opportunity, this account is unsatisfactory because it leaves unclear what falls *outside* the scope of social origins and thus remains as a basis for merit. The SC literature rarely offers

[2]Siblings' influence on one another may either increase or decrease similarity through processes such as role modeling and emulation on the one hand, and deidentification and niche-picking on the other (Jenkins and Dunn, 2009). Beenstock (2008, p. 325) claims that sibling interaction accounts for 'most of' the sibling correlation in an Israeli sample. The role of sibling interactions has been little studied in this literature, in part due to the inherent difficulties of identifying inter-sibling effects (Conley, 2008).

a direct answer. Rather, it is usually noted as a limitation that the SC does not fully capture the effects of family background. Discussing how to move from this lower bound towards a more complete measure, Björklund and Jäntti (2012, p. 468) survey the factors that ideally would be included. Among these are 'all [one's] (initial) genes' and 'all family and environmental influences.' But what else is there? This account leaves no conceptual room for the individual *as distinct from* their family background, and thus no basis for merit. Equality of opportunity is valorized for describing a world in which, for example, a hard-working individual can overcome the constraints imposed by family background. But the concept becomes incoherent if the characteristic of being hard-working itself necessarily falls within the all-consuming scope of 'family background.'

We argue instead that SCs make sense as a measure of inequality of opportunity only under an understanding of social origins as limited to common inheritance. Thus, there is conceptual room for merit, and merit on this account has its origins in individualizing factors. If individualizing factors have a relatively strong influence, the sibling correlation is relatively low and equality of opportunity relatively high. And vice versa where common inheritance plays the greater role. We do not propose that the SC is thereby a perfect measure of inequality of opportunity, but this perspective helps clarify its strengths and limitations. It is attractive because it expunges from merit all differences between families of origin, yet allows that individual differences in merit exist.

This raises the question of how far the division between common inheritance and individualising factors corresponds to that between (non-meritocratic) ascription and (meritocratic) achievement. The previous section makes clear that the correspondence is imperfect. We lack the space for a full discussion, but in the next section touch on the issue as it relates to genetics, and here briefly address one source of sibling differences that may appear non-meritocratic: differential parental investment. Insofar as this occurs *in response* to differences in ability between siblings rather than being idiosyncratic, it is consistent with SCs measuring inequality of opportunity. A rich literature studies whether—and which groups of—parents *compensate for* or instead *reinforce* sibling differences in ability (Fan and Porter, 2020; Frijters *et al.*, 2013; Grätz and Torche, 2016; Hsin, 2012). However it is agreed that differential investment occurs in response to observed signals. Reinforcement, which decreases sibling similarity, is the efficient strategy to enable the maximal realisation of potential (Becker and Tomes, 1976). By contrast, compensation (Behrman *et al.*, 1982), or equalization, increases sibling similarity and in doing so constrains the expression of innate potential, albeit equalising outcomes within families.

Differential investment by birth order or sex, however, would be a non-meritocratic source of sibling differences. More generally, this discussion highlights the value of sex-specific SCs. When brothers and sisters are both included, sex is treated as an individualizing factor: to the extent that it explains variation in the outcome, it will reduce the SC. Sex-specific SCs are thus more appropriate for measuring of inequality of opportunity, and indeed allow comparison across sexes.

Sibling correlations and genetically-sensitive variance decomposition

Genetically-informed designs such as the classical twin study allow a decomposition of variance into additive genetic (A), common environmental (C), and non-shared environmental (E) factors (Knopik *et al.*, 2016). This can inform debates on inequality of opportunity under the view that genetic differences represent differences in innate ability and thus merit or potential, whereas the common environment stands for ascribed factors such as parental wealth (Engzell and Tropf, 2019; Guo and Stearns, 2002; Nielsen, 2006). The literature linking behavioural genetics to social stratification has so far had relatively little to say about the interpretation of the non-shared environmental component, despite its large share in many contexts (Branigan *et al.*, 2013; Erola *et al.*, 2022).

The sibling correlation approach, by contrast, decomposes variance into two parts, $C + 0.5A$ (i.e. SC) and $E + 0.5A$ (i.e. $1 - SC$).[3] It is therefore an important limitation of the SC that it does not distinguish between genetic and environmental channels of transmission. Indeed, at the extremes, a given sibling correlation might be due entirely to C or entirely to A.

Nonetheless, we believe the two approaches are complementary. SC data is more easily collected and is available for a wider range of contexts (cf. Branigan *et al.*, 2013). Heritability as a measure of a context's equality of opportunity is not without limitations: in particular, the inherited traits that influence attainment may extend beyond ability and effort to non-merit characteristics such as skin color (Diewald *et al.*, 2015). Most importantly, behavioural genetics designs involve several much-debated assumptions that SCs do not (Felson, 2014; Fishkin, 2014; Wolfram and Morris, 2023).

How can the foregoing be reconciled with an interpretation of the SC as a measure of inequality of opportunity for outcomes subject to substantial genetic effects? Under the view that one's genome is simply *who one is* and is thus a legitimate basis for merit, the SC is limited as a measure of inequality of opportunity by the fact that it picks up the effect of the on average 50 per cent of genes shared with siblings. It follows from our account in the previous section that to interpret the SC as a measure of inequality of opportunity is instead

[3] In behavioural genetics, the sibling correlation is expected to be equivalent to the dizygotic twin correlation. Wolfram and Morris (2023) document divergence between the two and give an informative discussion.

to view that shared genetic component as a non-meritocratic factor, another (dis)advantage to happening to be born in a particular family—and to view the individual-specific genetic component as part of what makes one an individual distinct from one's family background, and thus a basis for merit. A corollary of this interpretation is that 'merit' (with respect to a particular outcome) is relative to one's siblings.

Sibling correlations and parent–child associations

Parent–child associations are very widely available and offer easily interpretable measures of intergenerational persistence. SCs however have important comparative advantages insofar as inequality of opportunity is at issue. Whereas parent–child associations describe the strength of association between child education and parent education, SCs describe the strength of association between child education and a much wider range of social origin factors. If we are interested in the extent to which one's background as a whole constrains or enables educational success, the more expansive concept is desirable (Hout, 2015). Parent–child associations in education mask variation in the outcome according to income, wealth, and myriad other parental factors, even among those with the same level of parental education. 'Parent'–child associations are also often father–child associations in practice, ignoring the role of maternal characteristics.

How much of the influence of social origins is missed by parent–child associations is made explicit by comparing the SC to the variance in child education accounted for by parental education. The SC can be decomposed into the sum of two components: the squared intergenerational correlation and other factors generating sibling similarity but uncorrelated with parental education (Solon, 1999).[4] Björklund and Jäntti (2020, p. 5) find that squared intergenerational correlations in education for Sweden and the US (.09 and .21) fall far below SCs from the same countries (.43 and .60 respectively), and adding further observables to the model does little to close the gap: '"other factors" are the most important ones that siblings share.'

SCs may also be more suited to comparative research than parent–child associations. The correlation between parental education and other family background characteristics that affect education is likely to exhibit substantial variation across time and place. One source of such variation would be the 'great deal of heterogeneity . . . across countries' in the wage returns to schooling (Trostel *et al.*, 2002, p. 2); that is, the link between parental education and parental earnings. This means that the extent to which parent–child associations capture the influence of social origins may vary considerably and in unclear ways across contexts, as suggested by the figures for Sweden and the US.

[4]The intergenerational correlation is the correlation between parent and child education.

Additionally, since the SC depends upon the educational distribution within one generation rather than two, it is without difficulties of interpretation that arise from mapping persistence over two generations, across which the meaning and distribution of education may change (and in different ways in different countries). For example, Hertz *et al.* (2008) reach differing conclusions about global change in educational persistence depending on whether the intergenerational coefficient or correlation is used.[5] The choice between the two also strikingly changes the rank order of countries. Our meta-analysis can therefore further contribute to a picture of country rankings of intergenerational persistence, where currently there is considerable uncertainty (Strömberg and Engzell, 2023).

Educational inequality of opportunity in comparative perspective and the role of income inequality

Comparative social mobility research has shed a great deal of light on factors underlying variation across time and space in the educational and occupational opportunity structure of societies (Breen and Müller, 2020; Bukodi and Goldthorpe, 2021; Smeeding *et al.*, 2011). At the same time, most of this work is based on intergenerational associations and thus paints a potentially incomplete picture. Comparative research into inequality of educational opportunity using SCs has been sparse so far. One study compared SCs among recent cohorts from across Europe and the US (Grätz *et al.*, 2021). Though the authors describe potentially relevant ways in which the countries under study differ, including their educational institutions and level of income inequality, with only six cases they do not attempt to link measured contextual factors to the size of the SC. Nonetheless they interpret certain results as running 'contrary to the Great Gatsby Curve hypothesis' (Grätz *et al.*, 2021, p. 1027). Sieben and de Graaf (2001, 2003) tested whether SCs varied systematically with indicators of modernisation (energy consumption per capita), individualization (divorce rate), and socialism (seats in parliament held by leftist parties). Whereas the effect of measured family background variables was negatively related to all these indicators, the total effect of social origins was not significantly related to any, raising the possibility that social origins taken as a whole exhibit a more invariant association with education than observables-based approaches would lead us to believe.

The Great Gatsby Curve (GGC) refers to the positive cross-national association between income inequality and intergenerational persistence, usually put

[5] Azam and Bhatt (2015) provide another illustration, in the context of India.

in terms of income or earnings. Theoretical accounts of the mechanisms underlying this relationship focus almost exclusively on how economic inequality creates educational inequality (Durlauf *et al.*, 2022; Ermisch *et al.*, 2012). However we are aware of only a few studies testing whether income inequality is cross-nationally associated with greater intergenerational persistence in education, none of which employ SCs (Blanden, 2013; Jerrim and Macmillan, 2015; Neidhöfer *et al.*, 2018). As Blanden (2013, p. 58) argues, evidence of an association between income inequality in childhood and educational persistence is in line with a causal interpretation of the GGC, since it is inconsistent with the idea that inequality and immobility 'tend to be generated by the same factors' (that is, are confounded) and therefore only correlate 'at the end of the process,' when the child generation are adults.

Income inequality's relation to the SC is likely to depend on the extent to which financial resources determine educational attainment. In terms of educational performance, or 'primary effects' (Boudon, 1974), greater income secures better health and nutrition and a home environment more conducive to learning (Breen *et al.*, 2009). As for educational decision-making conditional on performance, or 'secondary effects,' financial resources affect the perceived costs and benefits of continuing in education (Breen and Goldthorpe, 1997). Income is thus relevant to both primary and secondary effects, and higher income inequality indicates that it is more unevenly distributed. Under such circumstances we would expect a stronger effect of social origins on educational attainment (DiPrete, 2020; Durlauf *et al.*, 2022). Moreover, in high-inequality contexts the relative risk aversion mechanism of secondary effects (Breen and Goldthorpe, 1997) may be stronger, as the steeper income distribution implies a greater absolute loss from downward mobility.

On the other hand, for many of the cohorts in our study secondary education was free and tertiary education either free or subsidised to the point of widespread affordability, potentially nullifying the importance of income differences. Another possibility, since the SC captures influences beyond income, is that where incomes are relatively equal, high-status families develop and make greater use of cultural resources, social connections, and other substitutes for high income. The SC may then remain stable if advantaged parents are able to draw on alternative mechanisms through which to reproduce their position (Sieben and de Graaf, 2003).

The reduction of economic inequality may then be one policy option for weakening the link between origins and education. We aim to inform this debate by testing whether SCs correlate with cohort-specific indicators of income inequality at the national level. In common with other comparative macro-level research, this is necessarily a descriptive exercise. Rather than providing causal evidence, the benefits of our analysis lie in testing whether the hypothesized association is evident across a wide range of contexts, and in identifying where and when deviations appear.

Data and method

Our analysis has three main stages. The first comprises a systematic review. We undertook a literature search with the aim of identifying all studies which report at least one estimate of an SC in educational attainment. We extracted these estimates and other estimate-level information (e.g. sample size, dataset, model), constructing thereby a dataset of SC estimates nested in studies. The Supplementary Materials contain a list of the included studies and a reproduction package including the dataset of estimates we have assembled. Our descriptive statistics characterise this raw dataset, giving an overview of the literature.

Second, we use meta-regression to explore variation in the SC by country and (decadal) birth cohort.[6] We use random-effects models in all cases. In a preliminary step, because multiple estimates are often based on the same sample, we save one predicted estimate per sample. These predicted estimates come from a meta-regression of estimates on sibship type, model type, and sample fixed effects. At this stage we also exclude estimates from non-nationally representative samples, and those with a range of birth years so wide they cannot be assigned to a particular birth cohort.

Third, after using meta-analysis to pool estimates from different samples of the same country and cohort, we examine associations between these country-cohort estimates and indicators of income inequality.

The remainder of this section provides further detail underlying our analysis. Figure 1 provides a roadmap of our approach.

Literature search and inclusion criteria

Full details of our literature search and inclusion criteria are given in the Supplementary Materials. Here we state the main points. We included estimates of sibling correlations in educational attainment (*not* similar concepts such as test scores or enrolment), measured either in years of completed education or year-equivalents based on the usual time taken to attain a stated level. We only included estimates that refer to *brothers*, *sisters*, or to *all siblings* regardless of sex. We excluded estimates based on samples of individuals who either had not completed their education, or for whom being observed depended on residing in

[6] We focus on these results rather than meta-analysis per se because there is no meaningful 'one true' value (or distribution of values) to home in on; rather, sources of variation in the SC are of primary interest. For this reason, we do not report all the analyses and statistics one might typically see in a meta-analysis—but see Figure S2 in the Supplementary Materials for a (symmetrical) funnel plot and a brief discussion of publication bias.

Unit of analysis and *procedure*			Results	
Literature search		→	Figure S1*	Flowchart
↓				
All sibling correlation estimates meeting inclusion criteria ($N = 300$)		→	Figure S2*	Funnel plot
↓			Table 1	Descriptives
↓	*Exclude if non-nationally representative or sample cohort spans > 20 birth years*		Figure 2	Descriptives
↓			Figure S3*	Descriptives
Sibling correlation estimates ($N = 253$), including multiple estimates per sample				
↓				
↓	*RE meta-regression on sample (i.e. dataset–cohort), sibship type, and model type.*	→	Table 2	Regression output
↓			Table S1*	Estimate coverage by sibship type
↓	*Save predicted values for each sample, holding sibship type constant at* All siblings *and model type constant at* Pearson's r			
↓				
One estimate per (dataset–cohort) sample ($N = 94$)		→	Figure S4*	Forest plot
↓	↓			
↓	↓ RE meta-regression on country and cohort	→	Figure S5*	Coefficient plot
↓	↓			
↓	Predicted sibling correlations by country, holding cohort at 1960	→	Figure 3	Bar chart
↓			Figure S6*	Bar chart
↓			Figure S7*	Coefficient plot
↓	*RE meta-analysis within each country–cohort*			
↓				
Country–cohort predicted sibling correlations ($N = 80$)		→	Table 3	SCs by country and cohort
↓			Table S2*	SCs and income inequality data
↓				
↓	*RE meta-regressions on income inequality*	→	Figure 4	Bubble plot
			Table 4	Regression output
			Table 5	Margins on correlation scale

Figure 1: An overview of the analysis

Notes: *in Supplementary Materials. RE: Random effects. SC: Sibling correlation.

the parental home. From each study, we took the estimates that were the most disaggregated in terms of birth cohort.

Data extraction

For each estimate within a study, we extracted the following further information: *country*, *dataset*, *sibship type*, *sample size*, *birth years*, *model*, *nationally representative sample*, *singletons included*, and *adjusted for family size*. Here we provide a brief discussion of the listed factors which are not self-explanatory.

Model distinguishes different approaches to estimation of the sibling correlation: Pearson correlations, intraclass correlation coefficients (ICCs) from variance decomposition models, and (rarely) path coefficients from structural equation models. We collect information further distinguishing whether Pearson correlation estimates are *adjusted for family size*, and whether ICC estimates *include or exclude singletons* (only-children).[7] *Sample size* refers to the number of families within which the siblings are nested. This is to ensure comparability across SCs calculated from different approaches.

Data preparation

After describing the full set of SCs, we use meta-regression to explore variation in these estimates. To facilitate this, we make some adjustments to the data.

First, because we aim to make cross-national and cross-cohort comparisons, we exclude estimates from samples that are not nationally representative or whose birth years span more than 20 years. We assign estimates to decadal birth cohorts (e.g. 1950s) according to the midpoint of the birth years of the sample. Including such a cohort variable allows us to test for a cross-national time trend in the sibling correlation, as well as to purge our country coefficients of any confounding arising from such a trend.

Data preparation

After describing the full set of sibling correlations, we use meta-regression to explore variation in these estimates. To facilitate this, we make some adjustments to the data.

First, because we aim to make cross-national and cross-cohort comparisons, we exclude estimates from samples that are not nationally representative or whose birth years span more than 20 years. We assign estimates to decadal birth cohorts (e.g. 1950s) according to the midpoint of the birth years of the sample. Including such a cohort variable allows us to test for a cross-national

[7]When *sibling pairs* are weighted equally regardless of family size, large families are over-represented. Most studies recognize the issue and adjust so that each *family* is weighted equally. The implications of including singletons are discussed below.

time trend in the sibling correlation, as well as to purge our country coefficients of any confounding arising from such a trend.

Second, we follow established meta-analytic procedure (Borenstein *et al.*, 2009; Card, 2012) and apply the Fisher's z transformation (inverse hyperbolic tangent) to the extracted correlation coefficients r:

$$z = 0.5 \times \ln(\frac{1+r}{1-r}) \tag{2}$$

Whereas the variance of r depends on the value of r itself, this is not true of z. The standard error of z is simply a function of sample size n:

$$SE_z = \frac{1}{\sqrt{n-3}} \tag{3}$$

Meta-analytic techniques aggregate evidence, but it is important to ensure that evidence is not double-counted. This could happen if multiple studies estimate the same parameter for the same sample. Meta-analysis has its origins in the context of medical trials, where this appears not to be a widespread problem. In quantitative social science however, different studies inevitably draw on the same datasets.

Third, then, in order to yield a single comparable SC estimate for each sample of siblings, we run a meta-regression of estimates on sibship type, model type, and sample. We define samples as distinct combinations of dataset and decadal birth cohort. We save predicted values from this meta-regression, holding sibship type and model type constant at *All siblings* and *Pearson's r* respectively. This step also allows us to examine whether estimates systematically vary in magnitude by sibship or model type, net of sample fixed effects.

Above, we highlighted the value of sex-specific sibling correlations for measuring inequality of opportunity. Ideally our analysis would have focused on these. However, our aim is comparison across contexts, and we found all-siblings estimates for a far wider range of countries and cohorts, and sister correlations for particularly few (see Table S1 in the Supplementary Materials).

Meta-analysis and meta-regression

Meta-analysis is a technique for synthesizing existing evidence, allowing researchers to estimate a single overall effect size based on multiple studies estimating the same quantity (Borenstein *et al.*, 2009). Conceptually, our starting point is a random-effects meta-analysis, which assumes a normal distribution of true effect sizes around a grand mean μ. The deviation of each estimate z_i from μ has two components, ζ_i and ϵ_i:

$$z_i = \mu + \zeta_i + \epsilon_i \qquad \zeta_i \sim N(0, \tau^2) \text{ and } \epsilon_i \sim N(0, \sigma_i^2) \tag{4}$$

τ^2 is between-study variance in the true effect size (for exposition we here assume one estimate per study), while σ_i^2 is within-study variance arising from sampling error. The standard deviation σ_i equals SE_z as shown above. Meta-analysis calculates a weighted mean of the observed effect sizes, yielding an estimate of μ. For our purposes this overall mean is of only secondary interest.

A meta-regression is a regression of effect sizes on study-level covariates. Meta-regression is used to model variation in the true effect size due to differences between studies, net of differences due to sampling variation:

$$z_i = \beta_0 + \beta_1 x_1 + \zeta_i + \epsilon_i \tag{5}$$

β_0 is an intercept, x_1 a vector of study characteristics, and β_1 a vector of coefficients. Weighted least squares is used to estimate the coefficients. Meta-regression is often an extension to a study presenting a meta-analysis. In our case we focus on the results of our meta-regressions.

Details of further data and methods will be introduced below as they become relevant.

Results

Description of the sibling correlation literature

Table 1: Sibling correlations from the literature: basic descriptive information

	N	Min.	Mean (SD)	Max.
Studies	63			
Estimates	300	.13	.46 (.09)	.75
All siblings	124	.13	.45 (.10)	.69
Brothers	115	.24	.47 (.09)	.73
Sisters	61	.35	.48 (.06)	.75
Countries	23			
Datasets	59			

Note: Counts Germany, East Germany, West Germany, England, Scotland, and the United Kingdom as separate countries. For the meta-regressions, Germany and West Germany are combined, as are England and the United Kingdom. The title of this study refers to the 18 countries included in the meta-regressions.

We begin by characterizing the SCs identified in our search. As Table 1 shows, 300 estimates from 63 studies met our inclusion criteria. All-siblings, brother, and sister correlations were present in a ratio of approximately 2:2:1. We found estimates for 23 countries, based on 59 distinct datasets ranging from national registers to primary data from small surveys. Estimates range from

.13 (Czechoslovakia) to .75 (US), though these values are based on rather small samples (34 and 200 families respectively). The mean is .45 for all-siblings correlations, and slightly higher for brother (.47) and sister correlations (.48).

Figure 2 provides further detail. Panel A shows a steady growth in the number of SC studies over time. Panel B shows that early studies focused on brothers, but the number of all-siblings estimates has since caught up. There are still relatively few estimates of sister correlations, and in fact 33 of these 61 come from a single study (Wiborg and Hansen, 2018) and refer to Norway. Though the majority (63 %) of studies report just one or two estimates, a small group contain a large number. Panel C identifies six studies that together contain 157 estimates. These six include two which present estimates from Norwegian register data for many distinct cohorts, and four of the five previous studies to compare SCs cross-nationally.

Panels D and E of Figure 2 show which countries and datasets, respectively, are most represented in the literature. 31 studies have estimated SCs for US samples, by far the most. Next are the Nordic countries and Europe more widely. Apart from Australia, the non-European countries represented in the literature each appear in only one study. Notably, these are all (Indonesia, Israel, and Japan) excluded from our meta-regressions because the samples they refer to span a wide range of birth years. Besides register data, the PSID stands out as much-used. More recently-initiated household panel studies from other countries may be an untapped source of new SC data, since they largely adhere to the PSID's design. Panel F of Figure 2 shows the distribution of SCs: approximately normal, with 94 % of estimates falling within two standard deviations of the mean.

Figure S3 in the Supplementary Materials plots some additional descriptive information. As Panel A of Figure S3 shows, the sample sizes behind these estimates span several orders of magnitude, and are concentrated among the small and the very large. Panel B gives the distribution of the first year of birth for each estimate's sample. Coverage spans the late 19th century and most of the 20th, with younger cohorts yet to complete their education. The distribution of samples peaks in the 1950s and early 1960s. Panel C gives the length of these cohorts, that is, how long a span of birth years is included in each sample. This exhibits a long tail of studies in which nationally representative samples of adults were asked about their siblings' education. Usually however the estimates refer to more specific birth cohorts spanning under ten years.

Meta-analytic and meta-regression results

Before making comparisons across countries and cohorts, we regress estimates on sibship type, model type, and sample fixed effects. Using this model, we predict a single, comparable SC estimate for each sample. This also constitutes

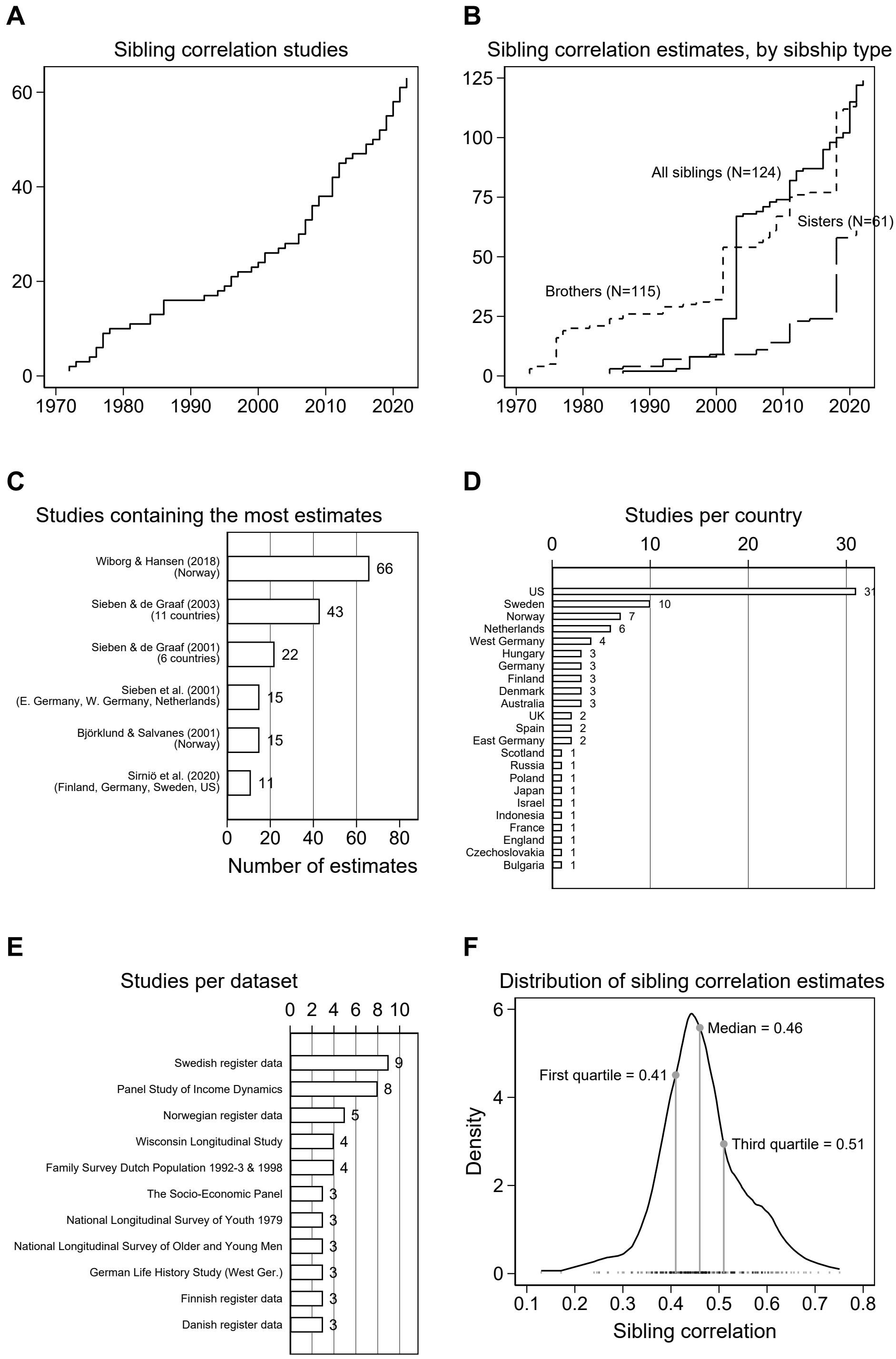

Figure 2: An overview of the sibling correlations literature. *Panel A*: cumulative frequency of sibling correlation studies over time. *Panel B*: cumulative frequency of sibling correlation estimates over time, by sibship type. *Panel C*: studies containing more than 10 estimates. *Panel D*: number of studies per country. *Panel E*: number of studies using each dataset. *Panel F*: kernel density plot of all sibling correlation estimates (Epanechnikov kernel).

Note to Panel E: Datasets used in one or two studies not shown.

Table 2: Random-effects meta-regression of sibling correlation estimates on sibship type, model type, and sample

	Coefficient (Fisher's z scale)	SE
Sibship type (*ref.* Brothers)		
All siblings	−0.014	0.009
Sisters	0.031***	0.006
Model type (*ref.* Pearson's r)		
Pearson's r, no adjustment for family size	−0.006	0.023
Variance decomposition, only-children excluded	0.014	0.013
Variance decomposition, only-children included	0.029	0.016
Structural equation model	−0.017	0.043
Model unclear	−0.007	0.017
Sample (dataset–cohort) fixed effects	included	
Observations (sibling correlation estimates)	253	

Note: $^{***}p < .001$.

a test of whether SC estimates vary systematically by sibship and model type. Coefficients from this model are shown in Table 2. Sister correlations are larger than those among brothers ($b = 0.031$) and those among all siblings ($b = 0.045$). The coefficients must be interpreted carefully, as they are on the Fisher's z scale and so vary across the range of correlation values; z is almost equivalent to r in the range 0 to .5, but notably diverges from r at $r \geq .5$. For the median estimate recorded in our literature review, .46, an increase of 0.031 would represent an increase of 0.024 on the correlation scale.

The three predominant approaches to estimation (*model type* in Table 2) are variance decomposition models either including ($N = 85$) or excluding (61) only-children, and Pearson correlations with adjustment for family size if applicable (68). Variance decompositions including only-children are expected to be larger, since only-children add between- but not within-family variance. This is reflected in their positive coefficient. Nonetheless, there are no significant differences among these three main approaches or the few estimates from structural equation models (5), but contrasts were significant between variance decompositions including only-children and both the 26 estimates whose model type we could not discern, and Pearson correlations without the family size adjustment (10).

We conclude that estimates are somewhat larger among sisters, estimates for all siblings are not significantly different from those among brothers, and the way in which a sibling correlation is estimated generally makes little difference. This is a reassuring note on which to predict a single estimate for each sample, holding sibship type at *all siblings* and model type at *Pearson's* r, while acknowledging that for many samples we in fact only observe correlations for brothers or sisters only, or estimated from other types of model. This yields estimates for 94 distinct samples. A forest plot in the Supplementary Materials (Figure S4) shows each estimate with a confidence interval and relative weight, plus the overall meta-analytic average of .46 (95 % CI .44, .47). The I^2 statistic of 90 % indicates that variation between these estimates is almost entirely due to between-estimate variance in the true effect size, rather than variance arising from sampling error.

Our focus is on modeling variation in these estimates rather than estimating an overall average which has no clear interpretation in this context. Figure S5 in the Supplementary Materials presents coefficients from a random-effects meta-regression of the 94 predicted correlations on country and birth cohort. The R-squared statistic indicates that these covariates explain 54.3 % of the variation between estimates. Estimates do not differ significantly according to birth cohort and thus produce no evidence of a cross-national time trend in the size of the SC.

The SC does however exhibit significant cross-national variation. The coefficients shown in Figure S5 are again on the Fisher's z scale; we first therefore discuss the rank ordering of countries and the significance of differences between them, rather than magnitude: Bulgaria and Spain exhibit the largest correlations, followed by France and then the US. The coefficient for the US does not differ significantly from that of the next few countries: Czechoslovakia, the Netherlands, Germany, Australia, Poland, Russia, and the United Kingdom. Continuing in descending order, Sweden, Hungary, Scotland, Norway, Finland, Denmark, and finally the former East Germany all exhibit significantly lower sibling correlations than the US (at the .05 level).

The random-effects model conservatively assumes that the true effect may vary even across different nationally representative samples of a given country and cohort, for instance due to subtle unmeasured differences in survey design and data quality. Under the alternative fixed-effects meta-regression, with its assumption that sampling error is the only source of differences between different estimates within each country and cohort, contrasts between the US on the one hand and the Netherlands, Germany, Australia, and the UK are all statistically significant at $p < .001$.

Though the US differs significantly from the Nordic countries under the random-effects assumption, it is worth emphasising that the random-effects model produces especially conservative standard errors in the case of the Nordic countries, where estimates are based exclusively on register data. The availability of such comprehensive and high-quality data makes the collection of further samples redundant. From the perspective of the random-effects model however, this is equivalent to sampling just one estimate from a normal distribution of true effects.

To give an indication of the magnitude of the differences between countries

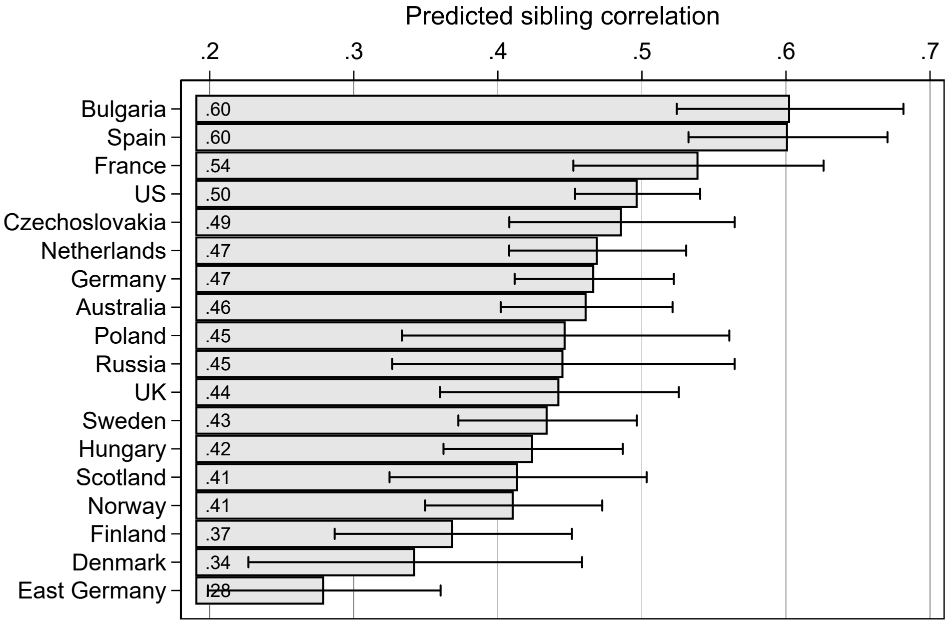

Figure 3: Predicted sibling correlations with 95 % confidence intervals, by country, from a random-effects meta-regression on country and cohort. Birth cohort held constant at 1960s.

Note: For the meta-regression analyses, England is coded as United Kingdom and West Germany as Germany.

on the correlation scale, we plot in Figure 3 predicted sibling correlations for each country, holding birth cohort at 1960. This yields predictions mostly in the range .4 to .5, with only Bulgaria, Spain, and France higher, and Finland, Denmark, and East Germany lower.

Income inequality

We next present meta-regressions of SCs on two different measures of income inequality. Each SC estimate in this case represents a single 'context,' or country–cohort combination, such as those born in France in the 1970s. To arrive at these country–cohort estimates ($N = 80$), we take, separately for each context, the weighted mean from a random-effects meta-analysis of all available samples of that context. These country–cohort estimates are shown in Table 3, which also indicates for which contexts each measure of income inequality is available. Table S2 in the Supplementary Materials additionally shows each context's value for each income inequality measure.

Table 3: Country–cohort meta-analytic sibling correlation estimates

	1900	1910	1920	1930	1940	1950	1960	1970	1980
Australia			*0.57*	*0.50*	*0.48*	***0.46***	***0.33***	***0.35***	
Bulgaria					0.44	**0.59**	**0.67**		
Czechoslovakia				0.12	**0.52**	**0.49**	**0.52**		
Denmark								***0.29***	***0.35***
East Germany			0.26	0.29	0.30	0.28	0.26		
Finland						**0.38**	**0.33**	***0.36***	
France						***0.51***		***0.53***	
Germany			*0.44*	0.48	*0.47*	*0.42*	***0.50***	***0.54***	***0.49***
Hungary	0.41	0.16	0.34	0.40	*0.51*	***0.49***	***0.54***		
Netherlands			*0.51*	*0.44*	*0.54*	***0.46***	***0.41***		
Norway				*0.40*	*0.42*	***0.42***	***0.41***	***0.40***	***0.39***
Poland						**0.41**	**0.47**		
Russia						*0.49*	*0.39*		
Scotland	0.55		0.39	0.39		0.42			
Spain			0.38	0.64	0.63	**0.57**	**0.68**		
Sweden				*0.45*	*0.45*	***0.41***	***0.40***	***0.44***	
United Kingdom	*0.48*		*0.52*	*0.44*				***0.35***	
United States	*0.58*	*0.57*	*0.56*	***0.49***	***0.47***	***0.48***	***0.51***	***0.40***	

Note: italics indicates the top decile share measure is available, **bold** indicates the Gini measure of income inequality is available. ***Bold and italic*** indicates both measures are available.

To measure income inequality, we first use Gini coefficients from the World Income Inequality Database (WIID) Companion dataset (Gradín, 2021a,b; UNU-

WIDER, 2022a).[8] In most cases this measure is only available from the 1960s onward. Second, we use top decile shares of pre-tax national income from the World Inequality Database (wid.world). This measure is available for more of our contexts and extends back to the early 20th century in some cases. In each case, we average all available values for the decade following the cohort's decade of birth, to capture conditions prevailing while sample members were in education.

Including all available country–cohort contexts, the association is weak and non-significant when using the Gini (p = .564, Figure 4). Driven by the early 20th century US, the association is significant when using top decile shares (p = .002). A cluster of high sibling correlation-low inequality points from Eastern Bloc countries raises the question of whether the hypothesis of a Great Gatsby Curve (GGC)-type association meaningfully applies in these contexts. Previous studies of the GGC which identified useable data for these countries have nevertheless excluded them, arguing that theorized GGC mechanisms depend upon democracy and a market economy (Andrews and Leigh, 2009; Jerrim and Macmillan, 2015).

Might it be possible to argue that the Eastern Bloc cases in fact represent a cluster of low-inequality countries which are within the scope of the GGC hypothesis but fail to support it? This argument does not seem strong. The literature attests a number of mechanisms *specific to the state-socialist context* by which advantaged parents were able to secure educational success for their children even under conditions of relative income inequality, Communist Party membership being an important example (Hanley, 2001; Hanley and McKeever, 1997; Shavit and Blossfeld, 1993).[9] Studies of inequality in postsocialist transitional societies emphasize the role of long-standing networks of clientelism (Nee and Cao, 2002). We therefore present results excluding these state-socialist contexts.

Results shown in Tables 4 and 5 extend our analysis and enhance its interpretability. Table 4 shows additional specifications including cohort and country fixed effects. For both measures, associations are similar albeit somewhat attenuated when analysing within-cohort cross-national variation. When analyzing within-country variation across cohorts, results diverge and the association remains evident with the top share measure—whether including or excluding the early 20th century US observations—but not with the Gini. This divergence would appear to result from the narrower range of cohorts and lower number

[8] The target income concept is net income per capita. Especially for richer countries, series gathered in the construction of the dataset mostly conform to this. Others 'have been adjusted to be consistent over time and across countries' (UNU-WIDER, 2022b, p. 3).

[9] Henderson *et al.* (2005) question the validity of the nominally low income inequality in the Eastern Bloc and, with respect to education, write that 'Communist decorations, parents' party membership, and references from the secretary of the local section of communist youth organizations all played a part in the entry process into higher education' (p. 402)

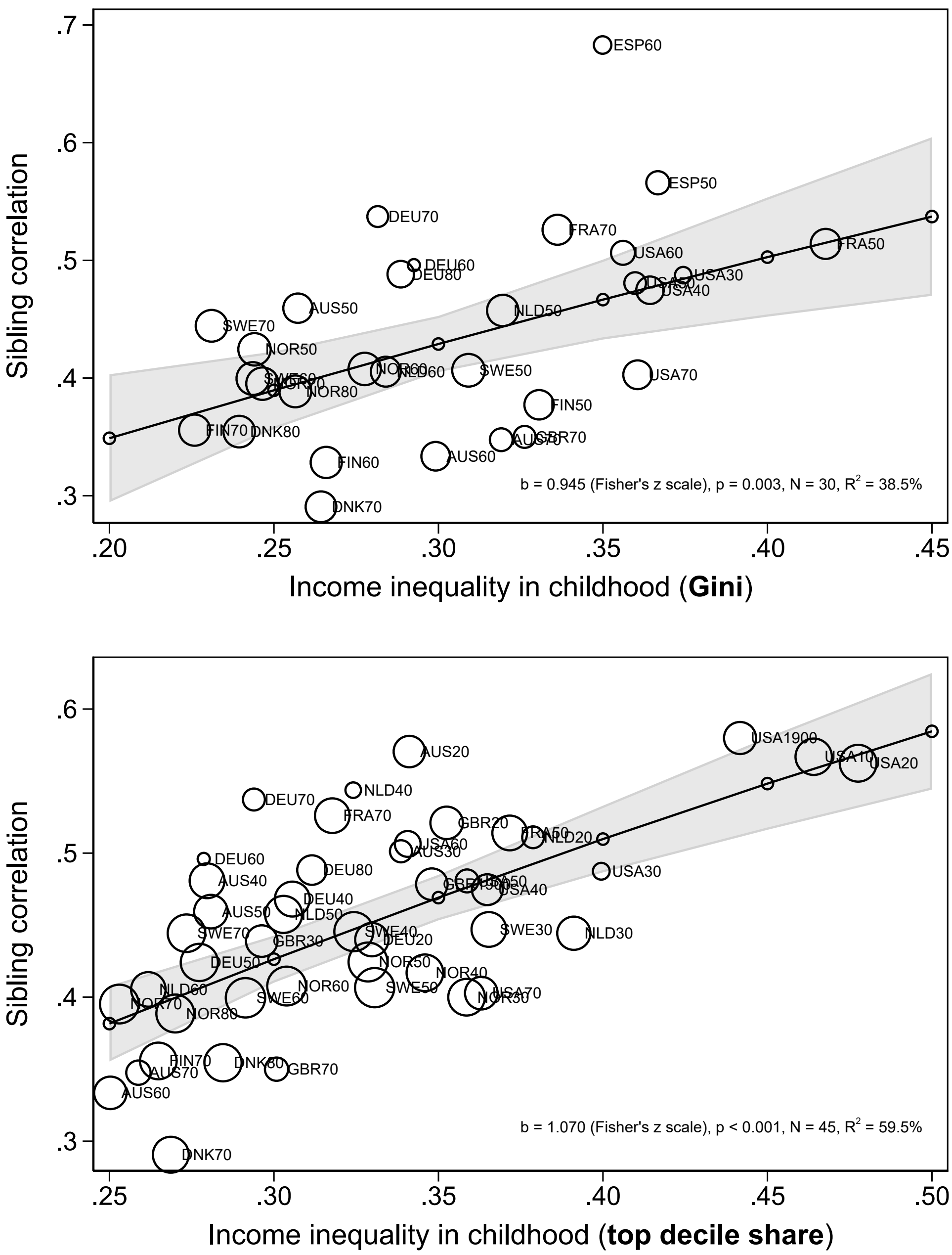

Figure 4: Country–cohort predicted sibling correlations and two measures of income inequality, with meta-regression lines and 95 % confidence intervals.

Note: The area of each observation's bubble is proportional to its weight in the meta-regression.

Table 4: Meta-regressions of country–cohort sibling correlation estimates on income inequality

	N			No	Cohort	Country	Cohort and
	Obs.	Cohorts	Countries	controls	FE	FE	country FE
Gini							
full sample	30	6	11	0.945**	0.914*	−0.256	−0.704
				(0.295)	(0.383)	(0.306)	(0.358)
both measures available	26	6	10	0.762**	0.649	−0.463	−0.900
				(0.258)	(0.323)	(0.357)	(0.371)
Top decile share							
full sample	45	9	10	1.07***	0.734**	0.874***	0.336
				(0.168)	(0.267)	(0.209)	(0.430)
excluding US 1900-1920	42	8	10	0.956**	0.724*	0.655*	0.288
				(0.254)	(0.342)	(0.242)	(0.466)
both measures available	26	6	10	1.173**	1.146*	0.377	−0.769
				(0.330)	(0.481)	(0.378)	(0.932)

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. FE: fixed effects. Coefficients are on the Fisher's z scale; see text.

Table 5: Meta-regressions of country–cohort sibling correlation estimates on income inequality: predictive margins on the correlation scale at selected values

		Predicted sibling correlation at Gini/top decile share of		Δ for .1 change from .25 to .35	
	N	.25 (A)	.35 (B)	(B) − (A)	SE of Δ
Gini					
no controls	30	0.390	0.467	0.077**	0.024
cohort FE	30	0.391	0.466	0.075**	0.031
country FE	30	0.420	0.399	−0.021	0.025
Top decile share					
no controls	45	0.382	0.469	0.087**	0.014
cohort FE	45	0.402	0.462	0.060**	0.022
country FE	45	0.391	0.462	0.071**	0.017

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. FE: fixed effects.

of observations overall for which the Gini measure is available. There are more than three observations for only Norway and the US, which each exhibit little change in the Gini across the period covered. The top decile share result is similarly non-significant and imprecisely estimated when restricting analysis to the Gini sample. Estimates become highly imprecise when including both cohort and country fixed effects. We thus find qualified support for an association between income inequality and the SC both cross-nationally within cohorts, and within countries across cohorts.

Because the meta-regressions are run on estimates transformed to the Fisher's z scale (as is the norm in meta-analysis), the slopes of the meta-regressions are slightly nonlinear on the original correlation scale. Figure 4 and Table 4 should therefore be interpreted in conjunction with Table 5, in which we present change in the predicted SC across a relevant .1 contrast in our contextual indicators. Focusing on the results shown in Figure 4, a .1 increase in the Gini from .25 to .35 is associated with a .077 increase in the SC. Moving from a top decile share of 25 % of pre-tax income to 35 % is associated with a .087 increase in the SC—comparable to the difference between the US and Norway presented in Figure 3.

Existing estimates of GGCs take countries as the unit of analysis and focus on a single birth cohort. Cohort fixed-effects are one way of accounting for dependence between multiple cohorts from the same country. As a robustness check we additionally re-run the meta-regressions from Figure 4 with the robust variance estimator of Hedges *et al.* (2010), using the user-written Stata command *robumeta* (Hedberg, 2014). The associations are almost identical in magnitude and remain statistically significant with both the Gini (p = .024) and top decile share measures (p = .006; p = .014 excluding the three early US cohorts).[10]

Discussion

The sibling correlation (SC) represents an estimate of the total effect of social origins, capturing the influence of all factors that increase sibling resemblance on an outcome. We have termed such factors, which include ethnicity and all aspects of parental socio-economic position, a sibling set's 'common inheritance,' and contrasted it with the 'individualizing factors' that generate sibling differences. The SC thus measures inequality of opportunity insofar as common inheritance corresponds to ascription and individualising factors to achievement.

SCs in education have been much-studied, but only rarely used in compar-

[10]For the p = .006 result (only), *robumeta* returns a warning that the p-value is untrustworthy due to low degrees of freedom (Tipton, 2015). This is due to the early US cohorts being outliers in terms of top decile share and reflects the relative paucity of observations at such high top decile share values.

ative research despite their attractive properties for this purpose. We aimed to exploit this opportunity by systematically searching the literature and collating all available estimates of SCs in educational attainment, using meta-regression to explore variability in the estimates, and examining how they vary across contexts marked by different levels of income inequality. Existing accounts of the link between economic inequality and intergenerational persistence have suggested educational inequality of opportunity plays an important intermediating role, but prior research has not explored whether the total effect of social origins on educational attainment systematically varies with income inequality.

Our two most important findings concern country differences in SCs and the strong positive association between the SC and income inequality.

The SC exhibits substantial cross-national variation. Of the (primarily European) countries for which we find estimates, Spain and Bulgaria stand out as having the largest correlations (.60; see Figure 3), while the former East Germany is at the lower extreme (.28), echoing the findings of Betthäuser's (2019) study of German reunification. Within the narrower range spanned by the majority of countries, France (.54) and the US (.50) are at the top end, and Norway (.41), Finland (.37), and Denmark (.34) at the lower end. It is important to understand the strengths and limitations of producing a single number for each country. These should be interpreted as reflecting average differences between countries, for cohorts born across the early to late-middle 20th century. This is necessarily an exercise in data smoothing.

These results nonetheless make an informative contribution to the comparative study of educational inequality. Prior country rankings have all been based on associations between parental education or occupation, and children's education (Beller and Hout, 2006; Hertz *et al.*, 2008; Jerrim and Macmillan, 2015; Pfeffer, 2008). They capture only part of the effect of social background on education—a part whose size as a portion of the total effect may vary cross-nationally. A point of wide interest as regards such rankings, and a popular topic more generally, is the position of the US relative to the Nordic countries (Andrade and Thomsen, 2021; Heckman and Landersø, 2022; Karlson and Birkelund, 2024). We find significantly lower SCs in Sweden, Norway, Finland, and Denmark than the US, with US sibling correlations on the order of .1—or 25 %—higher.[11][12] In terms of the rank ordering of this subset of countries, our results are thus broadly in agreement with Jerrim and Macmillan (2015) and the intergenerational correlation results of Hertz *et al.* (2008), but not—particularly

[11] This is also true when excluding SCs for cohorts born in the 1920s and earlier—cohorts for which we lack data on the Nordic countries, and for which the US sibling correlations are large. The only exception is that the difference with Sweden becomes non-significant (p = .072). Overall, this exclusion mainly affects Hungary, whose predicted correlation increases from .42 to .48. See Figure S6 in the Supplementary Materials.

[12] .1 and 25 % refer to .4 as an approximate average for the Nordic countries, and the US estimate of .5.

as regards Sweden and Norway—with those of Beller and Hout (2006), Pfeffer (2008), or Hertz *et al.*'s (2008) intergenerational coefficient results. It is beyond the scope of this study to explain differences among these other rankings, but future research should do so in the context of our results and, further, try to identify and comparatively study the unobserved factors captured in the sibling correlation that are masked in measures based on intergenerational associations. More data collection and analysis inclusive of all of the 'big four' dimensions of socio-economic status would be an important step in this direction, as would work using genetically-sensitive designs (Hällsten and Thaning, 2022).

We find larger SCs in educational attainment among cohorts growing up under conditions of greater income inequality. We are aware of no prior tests of this relationship, and the results provide support for theoretical accounts of the Great Gatsby Curve that emphasize education as a mediating pathway in the link between economic inequality and social immobility (DiPrete, 2020; Durlauf *et al.*, 2022). The mechanisms implied by these accounts are varied and multi-sited, from constraints on family investment, to neighborhood and school segregation, to the effects of inequality on the political process and support for public education. The policy implication of our finding is that greater equality of educational opportunity likely requires a reduction in economic inequality.

Great Gatsby Curve-type findings typically present a cross-sectional association. We harness observations of multiple cohorts to also test for a within-country association, and find qualified support for this. The finding is sensitive to the measure of income inequality used, though primarily because one is available for a wider range of contexts. Behind this overall finding, the data we have collated show that the US descends in both income inequality and the SC in three fairly distinct steps (1900–20, 1930, 1940–70), and Australia follows a similar trend.[13] Albeit somewhat noisily, the Netherlands, Sweden, and UK also tend in this direction across cohorts. Norway's SC is stable in the face of a declining top share of income, while Germany's exhibits trendless fluctuation against a stable top share. The association between income inequality and the SC is complex and contingent, but our results show a correlation between the two in aggregate.

We have three other noteworthy results. First, despite seeking all available SC estimates, in the end our sample is one of mostly mid-20th century cohorts from the Global North—specifically Europe, the US, and Australia. The main barrier to including existing estimates from Global South countries was sample inclusion depending on co-residence; education is likely to influence co-residence patterns among adult siblings.[14] The addition to existing surveys of questions

[13] Table 3 in the Supplementary Materials enables inspection of within-country change over time and with respect to the contextual measures.

[14] See Ahsan *et al.* (2022) for a study of sibling correlations in 53 developing countries. These estimates do not meet our inclusion criteria both because of the selection on coresident

asking about family members outside the household would be valuable.

Our income inequality results are particularly striking in this light: for the most part we are looking through a narrow window offering a view dominated by low-inequality contexts. We miss many high-inequality contexts from the early 20th century, whole other regions in the mid-to-late 20th century, and the widespread increase in income inequality from the 1980s onward. We would expect at least as strong an association to be evident if we had more estimates from these three broad sets of contexts.

Second, SCs between sisters are modestly higher on average than between brothers and between all siblings regardless of sex. This implies that even in the mid-to-late 20th century Global North where our estimates are concentrated, women's educational opportunities have tended to be more strongly constrained by social origin factors than men's. Since sex is a source of inter-sibling difference, sex-specific SCs are a preferable measure of inequality of opportunity. Figure S7 in the Supplementary Materials shows that the pattern of country differences changes little when brother-only and sister-only correlations are used. However, we found sister correlations for a relatively narrow range of contexts and therefore used (modelled) all-siblings correlations to accomplish our aim of comparison across a wide range of contexts. Future studies should report sex-specific SCs in addition to all-siblings ones.

Third, we find no overall differences between cohorts in the size of the SC. While this result is based on a meta-regression including country as a covariate, one should interpret it whilst keeping in mind the unbalanced nature of our panel (see Table 3). Data on earlier cohorts from more countries may have yielded a different conclusion with respect to the overall pattern of temporal change. We have used our limited space to focus on other questions, but this result and the data we have collated may be of interest to scholars of temporal change in the origins–education link. Further research should compare our estimates to those from in-depth country-specific studies based on intergenerational associations, which have reached conflicting conclusions (Breen *et al.*, 2009; Breen and Müller, 2020; Shavit and Blossfeld, 1993).

An important limitation is that the estimates we analyze refer to years of educational attainment and thus assume that all years of education are of equal standing. However, advantaged families seek qualitative as well as quantitative differentiation in their children's educational credentials (effectively maintained inequality, Lucas, 2001). Our analysis is blind to such qualitative inequality, though we would expect it to play the most pronounced role in countries with an elite stratum of higher education, such as the US, UK, and France. Nordic higher education is less stratified, suggesting that our main conclusion contrasting the latter with the US is only reinforced by this consideration. We also

status and because 'the sample consists of children of age 16–28' (M.S. Emran, personal communication, 5 July 2022). See also Emran *et al.* (2018).

note cautiously that that Pfeffer (2008, p. 549) undertakes a stability analysis incorporating differentiation into academic and vocational streams, and reports that 'These atalyses do not yield any evidence of systematic bias arising from the exclusive focus on the vertical dimension of the educational hierarchy.'

We have already discussed the nature of SCs at some length. Briefly, however, some further limitations should be noted. First, SCs are restricted to families with two or more children in most cases, albeit only-children can be included in a variance decomposition approach. Across the 20th century, only-children have tended to comprise no more than 12.5 % of each birth cohort in the countries we examine, Russia being an exception (Präg *et al.*, 2020). Differences between only-children and others in terms of social background and educational outcomes vary across contexts, but we are aware of no reason to believe the explanatory power of family background varies systematically between the two groups (Choi and Monden, 2024; Grätz *et al.*, 2021). Further, we have shown above that conditional on sibship type and dataset–cohort, estimates including only-children are not significantly different from those that exclude them, in the set of estimates we find. Nonetheless, this is a substantial subpopulation that most of our estimates exclude.

Second, as regards comparability, the need for data on multiple siblings entails a risk of studies diverging in their sample construction in potentially consequential ways. Our inclusion criteria help to mitigate this risk, and our meta-regression results suggest that the most prominent differences of approach are not systematically related to the size of the estimated correlation. Some studies imposed further restrictions which we did not measure and include, as most studies did not describe their sample construction to that level of detail. The main such concern surrounds restrictions on the age spacing of siblings, as siblings closer in age are likely to experience a more similar environment. However, an explicit comparison of results for closely- and widely-spaced siblings in Norway and the US 'does not find any substantial differences' (Björklund and Salvanes, 2011, p. 208). Another such concern is that differences may arise between designs where a parent reports their children's education and those where an individual reports their siblings' education. Sieben (2001) tests this possibility on a wide range of datasets and finds no significant differences. Finally, studies usually did not clarify whether half- or step-siblings were included. This likely reflects relatively low levels of family complexity across the contexts under study. However, family complexity has grown across societies in recent decades (Van Winkle, 2018; Van Winkle and Fasang, 2021), so future studies of younger cohorts will need to attend to this distinction.

Above, we compared SCs with heritability as a measure of equality of opportunity. One attraction of heritability is that, unlike the SC, it isolates the explanatory power of genes—within a given environmental context, it must be emphasized. This is an important consideration for educational attainment since heritability may widely exceed 50 % when assortative mating is taken into account (Wolfram and Morris, 2023). On the other hand, current molecular genetics estimates suggest far lower heritabilities (Fletcher *et al.*, 2023; Howe *et al.*, 2022), albeit methodological advances may in future reduce this missing heritability (Tropf *et al.*, 2017). We would sound a note of caution, however, as to the further claim that by disentangling common environment and genes, heritability automatically points to clear policy implications. 'Heritability does not imply immutability' (Plomin *et al.*, 2016, p. 4). It is by no means clear that genetic differences either cannot or should not be attenuated by educational policy (Engzell and Tropf, 2019; Fishkin, 2014; Manski, 2011).

We hope that the results reported here act as a spur for further research. A similar review of SCs in other stratification outcomes would be illuminating. We found comparisons of correlations for different outcomes within a given sample (Conley and Glauber, 2008) and tabulations of selected estimates of correlations in long-run earnings (Björklund and Jäntti, 2020; Schnitzlein, 2014), but no comprehensive account. Our dataset of SC estimates is available for others to use. We have also identified many contexts of interest for which we do not currently have such estimates. Estimates for cohorts that have recently completed their education will shed further light on the role of (recent increases in) income inequality. The more we can expand coverage beyond the contexts examined here, the richer will be our picture of the broad-scale institutional determinants of equality of opportunity.

References

Ahsan, Md. Nazmul, M. Shahe Emran, Hanchen Jiang, Qingyang Han, and Forhad Shilpi, 2022. 'Growing Up Together. Sibling Correlation, Parental Influence, and Intergenerational Educational Mobility in Developing Countries.' *SSRN* 4148659: 1–35. doi: 10.2139/ssrn.4148659.

Anderson, Lewis R., Patrick Präg, Evelina T. Akimova, and Christiaan Monden, 2024. 'Replication Materials to: The Total Effect of Social Origins on Educational Attainment. Meta-Analysis of Sibling Correlations From 18 Countries.' *Open Science Framework* doi: 10.17605/OSF.IO/XHCWR.

Andrade, Stefan B. and Jens-Peter Thomsen, 2021. 'Yes, Denmark Is a More Educationally Mobile Society than the United States. Rejoinder to Kristian Karlson.' *Sociological Science* 8(18): 359–370. doi: 10.15195/v8.a18.

Andrews, Dan and Andrew Leigh, 2009. 'More Inequality, Less Social Mobility.' *Applied Economics Letters* 16(15): 1489–1492. doi: 10.1080/13504850701720197.

Azam, Mehtabul and Vipul Bhatt, 2015. 'Like Father, Like Son? Intergenerational Educational Mobility in India.' *Demography* 52(6): 1929–1959. doi: 10.1007/s13524-015-0428-8.

Barclay, Kieron and Ken R. Smith, 2022. 'Birth Spacing and Health and Socioeconomic Outcomes Across the Life Course. Evidence From the Utah Population Database.' *Demography* 59(3): 1117–1142. doi: 10.1215/00703370-10015020.

Bates, John E., Alice C. Schermerhorn, and Isaac T. Petersen, 2012. 'Temperament and Parenting in Developmental Perspective.' In Zentner, Marcel and Rebecca L. Shiner, eds., *Handbook of Temperament*, pp. 425–441. New York: Guilford.

Becker, Gary S. and Nigel Tomes, 1976. 'Child Endowments and the Quantity and Quality of Children.' *Journal of Political Economy* 84(4, Part 2): S143–S162. doi: 10.1086/260536.

Beenstock, Michael, 2008. 'Deconstructing the Sibling Correlation. How Families Increase Inequality.' *Journal of Family and Economic Issues* 29(3): 325–345. doi: 10.1007/s10834-008-9114-y.

Behrman, Jere R., Robert A. Pollak, and Paul Taubman, 1982. 'Parental Preferences and Provision for Progeny.' *Journal of Political Economy* 90(1): 52–73. doi: 10.1086/261039.

Beller, Emily and Michael Hout, 2006. 'Intergenerational Social Mobility. The United States in Comparative Perspective.' *The Future of Children* 16(2): 19–36. doi: 10.1353/foc.2006.0012.

Betthäuser, Bastian A., 2019. 'The Effect of the Post-Socialist Transition on Inequality of Educational Opportunity. Evidence from German Unification.' *European Sociological Review* 35(4): 461–473. doi: 10.1093/esr/jcz012.

Björklund, Anders and Markus Jäntti, 2012. 'How Important Is Family Background for Labor-Economic Outcomes?' *Labor Economics* 19(4): 465–474. doi: 10.1016/j.labeco.2012.05.016.

——, 2020. 'Intergenerational Mobility, Intergenerational Effects, Sibling Correlations, and Equality of Opportunity. A Comparison of Four Approaches.' *Research in Social Stratification and Mobility* 70(100455): 1–11. doi: 10.1016/j.rssm.2019.100455.

Björklund, Anders and Kjell G. Salvanes, 2011. 'Education and Family Background. Mechanisms and Policies.' In Hanushek, Eric A., Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, volume 3, pp. 201–247. Amsterdam: Elsevier. doi: 10.1016/B978-0-444-53429-3.00003-X.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, 2005. 'The More the Merrier? The Effect of Family Size and Birth Order on Children's Education.' *Quarterly Journal of Economics* 120(2): 669–700. doi: 10.1093/qje/120.2.669.

Blanden, Jo, 2013. 'Cross-Country Rankings of Intergenerational Mobility. A Comparison of Approaches from Economics and Sociology.' *Journal of Economic Surveys* 27(1): 38–73. doi: 10.1111/j.1467-6419.2011.00690.x.

Borenstein, Michael, Larry V. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein, 2009. *Introduction to Meta-Analysis*. New York: Wiley. doi: 10.1002/9780470743386.

Boudon, Raymond, 1974. *Education, Opportunity, and Social Inequality*. New York: Wiley.

Branigan, Amelia R., Kenneth J. McCallum, and Jeremy Freese, 2013. 'Variation in the Heritability of Educational Attainment. An International Meta-Analysis.' *Social Forces* 92(1): 109–140. doi: 10.1093/sf/sot076.

Bredtmann, Julia and Nina Smith, 2018. 'Inequalities in Educational Outcomes. How Important Is the Family?' *Oxford Bulletin of Economics and Statistics* 80(6): 1117–1144. doi: 10.1111/obes.12258.

Breen, Richard and John H. Goldthorpe, 1997. 'Explaining Educational Differentials. Towards a Rational Action Theory.' *Rationality and Society* 9(3): 275–305. doi: 10.1177/104346397009003002.

Breen, Richard, Ruud Luijkx, Walter Müller, and Reinhard Pollak, 2009. 'Nonpersistent Inequality in Educational Attainment. Evidence from Eight European Countries.' *American Journal of Sociology* 114(5): 1475–1521. doi: 10.1086/595951.

——, 2010. 'Long-Term Trends in Educational Inequality in Europe. Class Inequalities and Gender Differences.' *European Sociological Review* 26(1): 31–48. doi: 10.1093/esr/jcp001.

Breen, Richard and Walter Müller, 2020. *Education and Intergenerational Social Mobility in Europe and the United States*. Stanford, CA: Stanford University Press. doi: 10.11126/stanford/9781503610163.001.0001.

Breinholt, Asta and Dalton Conley, 2023. 'Child-Driven Parenting. Differential Early Childhood Investment by Offspring Genotype.' *Social Forces* 102(1): 310–329. doi: 10.1093/sf/soac155.

Buchmann, Claudia, Thomas A. DiPrete, and Anne McDaniel, 2008. 'Gender Inequalities in Education.' *Annual Review of Sociology* 34: 319–337. doi: 10.1146/annurev.soc.34.040507.134719.

Bukodi, Erzsébet and John H. Goldthorpe, 2013. 'Decomposing 'Social Origins.' The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children.' *European Sociological Review* 29(5): 1024–1039. doi: 10.1093/esr/jcs079.

——, 2021. "Primary' Factors in Intergenerational Class Mobility in Europe. Results from the Application of a Topological Model.' *European Sociological Review* 37(1): 1–17. doi: 10.1093/esr/jcaa028.

Card, Noel A., 2012. *Applied Meta-Analysis for Social Science Research*. New York: Guilford.

Chan, Tak Wing, Morag Henderson, and Rachel Stuchbury, 2019. 'Family Size and Educational Attainment in England and Wales.' *Population Studies* 73(2): 165–178. doi: 10.1080/00324728.2019.1577479.

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang, 2017. 'The Fading American Dream. Trends in Absolute Income Mobility since 1940.' *Science* 356(6336): 398–406. doi: 10.1126/science.aal4617.

Choi, Seongsoo and Christiaan Monden, 2024. 'Where It Matters to Be the Only One. School Performance Outcomes of Only-Children Across 31 Countries.' *SocArXiv* pp. 1–64. doi: 10.31235/osf.io/kc6x5.

Conley, Dalton, 2008. 'Bringing Sibling Differences in. Enlarging Our Understanding of the Transmission of Advantage in Families.' In Lareau, Annette and Dalton Conley, eds., *Social Class. How Does It Work?*, pp. 179–200. New York: Russell Sage.

Conley, Dalton and Rebecca Glauber, 2005. 'Sibling Similarity and Difference in Socio-Economic Status: Life Course and Family Resource Effects.' *NBER Working Paper* 11320: 1–46. doi: 10.3386/w11320.

——, 2008. 'All in the Family? Family Composition, Resources, and Sibling Similarity in Socioeconomic Status.' *Research in Social Stratification and Mobility* 26(4): 297–306. doi: 10.1016/j.rssm.2008.08.003.

Corak, Miles, 2013. 'Inequality from Generation to Generation. The United States in Comparison.' In Rycroft, Robert S., ed., *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, volume 1, pp. 107–123. Santa Barbara, CA: ABC-CLIO.

Diewald, Martin, Tina Baier, Wiebke Schulz, and Reinhard Schunck, 2015. 'Status Attainment and Social Mobility.' *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 67(Supplement 1): 371–395. doi: 10.1007/s11577-015-0317-6.

DiPrete, Thomas A., 2020. 'The Impact of Inequality on Intergenerational Mobility.' *Annual Review of Sociology* 46: 379–398. doi: 10.1146/annurev-soc-121919-054814.

Durlauf, Steven N., Andros Kourtellos, and Chih Ming Tan, 2022. 'The Great Gatsby Curve.' *Annual Review of Economics* 14: 571–605. doi: 10.1146/annurev-economics-082321-122703.

Emran, M. Shahe, William Greene, and Forhad Shilpi, 2018. 'When Measure Matters. Coresidency, Truncation Bias, and Intergenerational Mobility in Developing Countries.' *Journal of Human Resources* 53(3): 589–607. doi: 10.3368/jhr.53.3.0216-7737R1.

Engzell, Per and Felix C. Tropf, 2019. 'Heritability of Education Rises with Intergenerational Mobility.' *Proceedings of the National Academy of Sciences* 116(51): 25386–25388. doi: 10.1073/pnas.1912998116.

Ermisch, John, Markus Jäntti, Timothy M. Smeeding, and James A. Wilson, 2012. 'Advantage in Comparative Perspective.' In Ermisch, John, Markus Jäntti, and Timothy M. Smeeding, eds., *From Parents to Children. The Intergenerational Transmission of Advantage*, pp. 3–31. New York: Russell Sage.

Erola, Jani, Hannu Lehti, Tina Baier, and Aleksi Karhula, 2022. 'Socioeconomic Background and Gene–Environment Interplay in Social Stratification across the Early Life Course.' *European Sociological Review* 38(1): 1–17. doi: 10.1093/esr/jcab026.

Fan, Wei and Catherine Porter, 2020. 'Reinforcement or Compensation? Parental Responses to Children's Revealed Human Capital Levels.' *Journal of Population Economics* 33(1): 233–270. doi: 10.1007/s00148-019-00752-7.

Felson, Jacob, 2014. 'What Can We Learn From Twin Studies? A Comprehensive Evaluation of the Equal Environments Assumption.' *Social Science Research* 43: 184–199. doi: 10.1016/j.ssresearch.2013.10.004.

Fishkin, Joseph, 2014. *Bottlenecks. A New Theory of Equal Opportunity.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199812141.001.0001.

Fletcher, Jason M., Qiongshi Lu, Bhashkar Mazumder, and Jie Song, 2023. 'Understanding Sibling Correlations in Education. Molecular Genetics and Family Background.' *IZA Discussion Paper* 15862: 1–19.

Frijters, Paul, David W. Johnston, Grace Lordan, and Michael A. Shields, 2013. 'Exploring the Relationship between Macroeconomic Conditions and Problem Drinking as Captured by Google Searches in the US.' *Social Science & Medicine* 84: 61–68. doi: 10.1016/j.socscimed.2013.01.028.

Gradín, Carlos, 2021a. 'WIID Companion May 2021. Data Selection.' *WIDER Technical Note* 7/2021: 1–17. doi: 10.35188/UNU-WIDER/WTN/2021-7.

——, 2021b. 'WIID Companion May 2021. Integrated and Standardized Series.' *WIDER Technical Note* 8/2021: 1–35. doi: 10.35188/UNU-WIDER/WTN/2021-8.

Grätz, Michael, 2018. 'Competition in the Family. Inequality between Siblings and the Intergenerational Transmission of Educational Advantage.' *Sociological Science* 5(11): 246–269. doi: 10.15195/v5.a11.

Grätz, Michael, Kieron J. Barclay, Øyvind Wiborg, Torkild H. Lyngstad, Aleksi Karhula, Jani Erola, Patrick Präg, Thomas Laidley, and Dalton Conley, 2021. 'Sibling Similarity in Education Across and Within Societies.' *Demography* 58(3): 1011–1037. doi: 10.1215/00703370-916402.

Grätz, Michael and Florencia Torche, 2016. 'Compensation or Reinforcement? The Stratification of Parental Responses to Children's Early Ability.' *Demography* 53(6): 1883–1904. doi: 10.1007/s13524-016-0527-1.

Grätz, Michael and Øyvind N. Wiborg, 2024. 'Parental Ages and the Intergenerational Transmission of Education. Evidence From Germany, Norway, and the United States.' *European Societies* 26(5). doi: 10.1080/14616696.2024.2310011.

Guo, Guang and Elizabeth Stearns, 2002. 'The Social Influences on the Realization of Genetic Potential for Intellectual Development.' *Social Forces* 80(3): 881–910. doi: 10.1353/sof.2002.0007.

Halliday, Timothy J. and Bhashkar Mazumder, 2017. 'An Analysis of Sibling Correlations in Health using Latent Variable Models.' *Health Economics* 26(12): e108–e125. doi: 10.1002/hec.3483.

Hanley, Eric, 2001. 'Centrally Administered Mobility Reconsidered. The Political Dimension of Educational Stratification in State-Socialist Czechoslovakia.' *Sociology of Education* 74(1): 25–43. doi: 10.2307/2673143.

Hanley, Eric and Matthew McKeever, 1997. 'The Persistence of Educational Inequalities in State-Socialist Hungary: Trajectory-Maintenance versus Counterselection.' *Sociology of Education* 70(1): 1–18. doi: 10.2307/2673189.

Healey, Matthew D. and Bruce J. Ellis, 2007. 'Birth Order, Conscientiousness, and Openness to Experience. Tests of the Family-Niche Model of Personality Using a Within-Family Methodology.' *Evolution and Human Behavior* 28(1): 55–59. doi: 10.1016/j.evolhumbehav.2006.05.003.

Heckman, James and Rasmus Landersø, 2022. 'Lessons for Americans From Denmark About Inequality and Social Mobility.' *Labor Economics* 77(101999): 1–14. doi: 10.1016/j.labeco.2021.101999.

Hedberg, Eric Christopher, 2014. 'robumeta: Stata Module to Perform Robust Variance Estimation in Meta-Regression With Dependent Effect Size Estimates.' *Statistical Software Components* S457219.

Hedges, Larry V., Elizabeth Tipton, and Matthew C. Johnson, 2010. 'Robust Variance Estimation in Meta-Regression With Dependent Effect Size Estimates.' *Research Synthesis Methods* 1(1): 39–65. doi: 10.1002/jrsm.5.

Henderson, David R., Robert M. McNab, and Tamás Rózsás, 2005. 'The Hidden Inequality in Socialism.' *The Independent Review* 9(3): 389–412. doi: 10.2307/24562283.

Hernández-Alava, Mónica and Gurleen Popli, 2017. 'Children's Development and Parental Input. Evidence From the UK Millennium Cohort Study.' *Demography* 54(2): 485–511. doi: 10.1007/s13524-017-0554-6.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina, 2008. 'The Inheritance of Educational Inequality. International Comparisons and Fifty-Year Trends.' *BE Journal of Economic Analysis and Policy* 7(2): 1–46. doi: 10.2202/1935-1682.1775.

Hällsten, Martin and Max Thaning, 2022. 'Wealth as One of the "Big Four" SES Dimensions in Intergenerational Transmissions.' *Social Forces* 100(4): 1533–1560. doi: 10.1093/sf/soab080.

Hout, Michael, 2015. 'A Summary of What We Know about Social Mobility.' *Annals of the American Academy of Political and Social Science* 657(1): 27–36. doi: 10.1177/0002716214547174.

——, 2018. 'Americans' Occupational Status Reflects the Status of Both of Their Parents.' *Proceedings of the National Academy of Sciences* 115(38): 9527–9532. doi: 10.1073/pnas.1802508115.

Howe, Laurence J., Michel G. Nivard, Tim T. Morris, Ailin F. Hansen, Humaira Rasheed, Yoonsu Cho, Geetha Chittoor, Rafael Ahlskog, Penelope A. Lind, Teemu Palviainen, Matthijs D. van der Zee, Rosa Cheesman, Massimo Mangino, Yunzhang Wang, Shuai Li, Lucija Klaric, Scott M. Ratliff, Lawrence F. Bielak, Marianne Nygaard, Alexandros Giannelis, Emily A. Willoughby, Chandra A. Reynolds, Jared V. Balbona, Ole A. Andreassen, Helga Ask, Aris Baras, Christopher R. Bauer, Dorret I. Boomsma, Archie Campbell, Harry Campbell, Zhengming Chen, Paraskevi Christofidou, Elizabeth Corfield, Christina C. Dahm, Deepika R. Dokuru, Luke M. Evans, Eco J. C. de Geus, Sudheer Giddaluru, Scott D. Gordon, K. Paige Harden, W. David Hill, Amanda Hughes, Shona M. Kerr, Yongkang Kim, Hyeokmoon Kweon, Antti Latvala, Deborah A. Lawlor, Liming Li, Kuang Lin, Per Magnus, Patrik K. E. Magnusson, Travis T. Mallard, Pekka Martikainen, Melinda C. Mills, Pål Rasmus Njølstad, John D. Overton, Nancy L. Pedersen, David J. Porteous, Jeffrey Reid, Karri Silventoinen, Melissa C. Southey, Camilla Stoltenberg, Elliot M. Tucker-Drob, Margaret J. Wright, Social Science Genetic Association Consortium, Within Family Consortium, John K. Hewitt, Matthew C. Keller, Michael C. Stallings, James J. Lee, Kaare Christensen, Sharon L. R. Kardia, Patricia A. Peyser, Jennifer A. Smith, James F. Wilson, John L. Hopper, Sara Hägg, Tim D. Spector, Jean-Baptiste Pingault, Robert Plomin, Alexandra Havdahl, Meike Bartels, Nicholas G. Martin, Sven Oskarsson, Anne E. Justice, Iona Y. Millwood, Kristian Hveem, Øyvind Naess, Cristen J. Willer, Bjørn Olav Åsvold, Philipp D. Koellinger, Jaakko Kaprio, Sarah E. Medland, Robin G. Walters, Daniel J. Benjamin, Patrick Turley, David M. Evans, George Davey Smith, Caroline Hayward, Ben Brumpton, Gibran Hemani, and Neil M. Davies, 2022. 'Within-Sibship Genome-Wide Association Analyses Decrease Bias in Estimates of Direct Genetic Effects.' *Nature Genetics* 54(5): 581–592. doi: 10.1038/s41588-022-01062-7.

Hsin, Amy, 2012. 'Is Biology Destiny? Birth Weight and Differential Parental Treatment.' *Demography* 49(4): 1385–1405. doi: 10.1007/s13524-012-0123-y.

Huh, David, Jennifer Tristan, Emily Wade, and Eric Stice, 2006. 'Does Problem Behavior Elicit Poor Parenting? A Prospective Study of Adolescent Girls.' *Journal of Adolescent Research* 21(2): 185–204. doi: 10.1177/0743558405285462.

Jenkins, Jennifer and Judy Dunn, 2009. 'Siblings Within Families. Levels of Analysis and Patterns of Influence.' *New Directions for Child and Adolescent Development* 2009(126): 79–93. doi: 10.1002/cd.258.

Jerrim, John and Lindsey Macmillan, 2015. 'Income Inequality, Intergenerational Mobility, and the Great Gatsby Curve. Is Education the Key?' *Social Forces* 94(2): 505–533. doi: 10.1093/sf/sov075.

Karhula, Aleksi, Jani Erola, Marcel Raab, and Anette Fasang, 2019. 'Destination as a Process. Sibling Similarity in Early Socioeconomic Trajectories.' *Advances in Life Course Research* 40: 85–98. doi: 10.1016/j.alcr.2019.04.015.

Karlson, Kristian B. and Jesper F. Birkelund, 2024. 'Origins of Attainment. Do Brother Correlations in Occupational Status and Income Overlap?' *European Sociological Review* 40(3): 379–389. doi: 10.1093/esr/jcad030.

Kiff, Cara J., Liliana J. Lengua, and Maureen Zalewski, 2011. 'Nature and Nurturing. Parenting in the Context of Child Temperament.' *Clinical Child and Family Psychology Review* 14(3): 251–301. doi: 10.1007/s10567-011-0093-4.

Knopik, Valerie S., Jenae M. Neiderhiser, John C. DeFries, and Robert Plomin, 2016. *Behavioral Genetics.* New York: Worth, 7th edition.

Lareau, Annette, 2011. *Unequal Childhoods. Class, Race, and Family Life.* Berkeley, CA: University of California Press, 2nd edition.

Lillehagen, Mats and Martin Arstad Isungset, 2020. 'New Partner, New Order? Multipartnered Fertility and Birth Order Effects on Educational Achievement.' *Demography* 57(5): 1625–1646. doi: 10.1007/s13524-020-00905-4.

Lucas, Samuel, 2001. 'Effectively Maintained Inequality. Education Transitions, Track Mobility, and Social Background Effects.' *American Journal of Sociology* 106(6): 1642–1690. doi: 10.1086/321300.

Lundberg, Ian, 2020. 'Does Opportunity Skip Generations? Reassessing Evidence From Sibling and Cousin Correlations.' *Demography* 57(4): 1193–1213. doi: 10.1007/s13524-020-00880-w.

Maccoby, Eleanor E., 2000. 'Parenting and its Effects on Children. On Reading and Misreading Behavior Genetics.' *Annual Review of Psychology* 51: 1–27. doi: 10.1146/annurev.psych.51.1.1.

Manski, Charles F., 2011. 'Genes, Eyeglasses, and Social Policy.' *Journal of Economic Perspectives* 25(4): 83–94. doi: 10.1257/jep.25.4.83.

Nee, Victor and Yang Cao, 2002. 'Postsocialist Inequalities. The Causes of Continuity and Discontinuity.' *Research in Social Stratification and Mobility* 19: 3–39. doi: 10.1016/S0276-5624(02)80035-X.

Neidhöfer, Guido, Joaquín Serrano, and Leonardo Gasparini, 2018. 'Educational Inequality and Intergenerational Mobility in Latin America. A New Database.' *Journal of Development Economics* 134: 329–349. doi: 10.1016/j.jdeveco.2018.05.016.

Nielsen, François, 2006. 'Achievement and Ascription in Educational Attainment. Genetic and Environmental Influences on Adolescent Schooling.' *Social Forces* 85(1): 193–216. doi: 10.1353/sof.2006.0135.

Pfeffer, Fabian T., 2008. 'Persistent Inequality in Educational Attainment and Its Institutional Context.' *European Sociological Review* 24(5): 543–565. doi: 10.1093/esr/jcn026.

Plomin, Robert, 2011. 'Why Are Children in the Same Family So Different? Non-shared Environment Three Decades Later.' *International Journal of Epidemiology* 40(3): 582–592. doi: 10.1093/ije/dyq144.

Plomin, Robert, John C. DeFries, Valerie S. Knopik, and Jenae M. Neiderhiser, 2016. 'Top 10 Replicated Findings From Behavioral Genetics.' *Perspectives on Psychological Science* 11(1): 3–23. doi: 10.1177/1745691615617439.

Präg, Patrick, Seongsoo Choi, and Christiaan Monden, 2020. 'The Sibsize Revolution in International Context. Declining Social Disparities in the Number of Siblings in 26 Countries.' *Demographic Research* 43(17): 461–500. doi: 10.4054/DemRes.2020.43.17.

Przeworski, Adam and Henry Teune, 1970. *The Logic of Comparative Social Inquiry*. New York: Wiley.

Putnam, Robert D., 2015. *Our Kids. The American Dream in Crisis*. New York: Simon & Schuster.

Roemer, Jon E., 1998. *Equality of Opportunity*. Cambridge, MA: Harvard University Press.

Schnitzlein, Daniel 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. doi: 10.1007/s00148-013-0468-6.

Shavit, Yossi and Hans-Peter Blossfeld, 1993. *Persistent Inequality. Changing Educational Attainment in Thirteen Countries*. Boulder, CO: Westview.

Sieben, Inge, 2001. *Sibling Similarities and Social Stratification. The Impact of Family Background across Countries and Cohorts*. Nijmegen: ICS/Radboud University.

Sieben, Inge and Paul M. de Graaf, 2001. 'Testing the Modernization Hypothesis and the Socialist Ideology Hypothesis. A Comparative Sibling Analysis of Educational Attainment and Occupational Status.' *British Journal of Sociology* 52(3): 441–467. doi: 10.1080/00071310120071133.

——, 2003. 'The Total Impact of the Family on Educational Attainment. A Comparative Sibling Analysis.' *European Societies* 5(1): 33–68. doi: 10.1080/1461669032000057668a.

Sieben, Inge, Johannes Huinink, and Paul M. de Graaf, 2001. 'Family Background and Sibling Resemblance in Educational Attainment. Trends in the Former FRG, the Former GDR, and the Netherlands.' *European Sociological Review* 17(4): 401–430. doi: 10.1093/esr/17.4.401.

Sirniö, Outi, Hannu Lehti, Michael Grätz, Kieron Barclay, and Jani Erola, 2020. 'The Pattern of Educational Inequality. The Contribution of Family Background on Levels of Education Over Time and Across Four Countries.' doi: 10.31235/osf.io/nupfs.

Smeeding, Timothy M., Robert Erikson, and Markus Jäntti, 2011. *Persistence, Privilege, and Parenting. The Comparative Study of Intergenerational Mobility.* New York: Russell Sage.

Solon, Gary, 1999. 'Intergenerational Mobility in the Labor Market.' In Ashenfelter, Orley C. and David Card, eds., *Handbook of Labor Economics*, volume 3A, pp. 1761–1800. Amsterdam: Elsevier. doi: 10.1016/S1573-4463(99)03010-2.

Sorokin, Pitrim A., 1927. *Social Mobility.* New York: Harper & Row.

Strömberg, Ely and Per Engzell, 2023. 'How Robust Are Country Rankings in Educational Mobility?' *SocArXiv* pp. 1–34. doi: 10.31235/osf.io/9zkc2.

Sulloway, Frank J., 1996. *Born to Rebel. Birth Order, Family Dynamics, and Creative Lives.* New York: Pantheon.

Tipton, Elizabeth, 2015. 'Small Sample Adjustments for Robust Variance Estimation With Meta-Regression.' *Psychological Methods* 20(3): 375–393. doi: 10.1037/met0000011.

Tropf, Felix C., S. Hong Lee, Renske M. Verweij, Gert Stulp, Peter J. van der Most, Ronald de Vlaming, Andrew Bakshi, Daniel A. Briley, Charles Rahal, Robert Hellpap, Anastasia N. Iliadou, Tõnu Esko, Andres Metspalu, Sarah E. Medland, Nicholas G. Martin, Nicola Barban, Harold Snieder, Matthew R. Robinson, and Melinda C. Mills, 2017. 'Hidden Heritability Due to Heterogeneity Across Seven Populations.' *Nature Human Behavior* 1(10): 757–765. doi: 10.1038/s41562-017-0195-1.

Trostel, Philip, Ian Walker, and Paul Woolley, 2002. 'Estimates of the Economic Return to Schooling for 28 Countries.' *Labor Economics* 9(1): 1–16. doi: 10.1016/S0927-5371(01)00052-5.

UNU-WIDER, 2022a. *World Income Inequality Database (WIID) Companion Dataset*. Helsinki: United Nations University, World Institute for Development Economics Research. doi: 10.35188/UNU-WIDER/WIIDcomp-300622.

——, 2022b. *World Income Inequality Database (WIID) Companion Dataset User Guide*. Helsinki: United Nations University, World Institute for Development Economics Research.

Van Winkle, Zachary, 2018. 'Family Trajectories Across Time and Space. Increasing Complexity in Family Life Courses in Europe?' *Demography* 55(1): 135–164. doi: 10.1007/s13524-017-0628-5.

Van Winkle, Zachary and Anette Fasang, 2021. 'The Complexity of Employment and Family Life Courses Across 20th Century Europe. More Evidence for Larger Cross-National Differences but Little Change Across 1916–1966 Birth Cohorts.' *Demographic Research* 44(32): 775–810. doi: 10.4054/DemRes.2021.44.32.

Wiborg, Øyvind N. and Marianne N. Hansen, 2018. 'The Scandinavian Model During Increasing Inequality. Recent Trends in Educational Attainment, Earnings, and Wealth among Norwegian Siblings.' *Research in Social Stratification and Mobility* 56: 53–63. doi: 10.1016/j.rssm.2018.06.006.

Wolfram, Tobias and Damien Morris, 2023. 'Conventional Twin Studies Overestimate the Environmental Differences Between Families Relevant to Educational Attainment.' *NPJ Science of Learning* 8(1): 24. doi: 10.1038/s41539-023-00173-y.

Supplementary Materials

Contents

A Link to reproduction package

Anderson *et al.* (2024) provides access to materials to reproduce our analysis:

- Stata code.
- Our dataset of estimates collected from the literature.
- Secondary datasets containing measures of income inequality.

B Details of literature search and full inclusion criteria

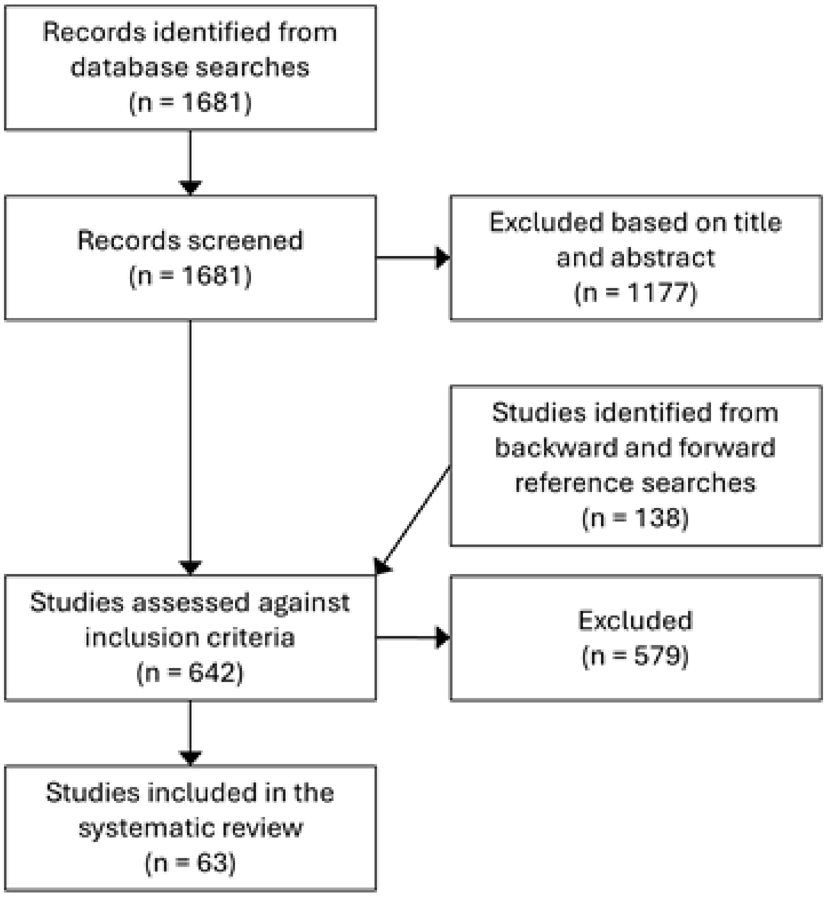

Figure S1: Flow diagram detailing the literature search and inclusion and exclusion of studies

We first searched the Web of Science and Sociological Abstracts databases. The only restriction imposed at this stage was that implied by only using

English-language search terms.[15] Our searches took the form 'E and SC', including every combination of the following for E, S, and C respectively:

- educational attainment, education, educational achievement, income, occupation, social class, socio-economic status;
- sibling, brother, sister, fraternal, sororal;
- correlation, similarity, dissimilarity, resemblance.

Figure S1 details the steps of our search. We used several additional strategies to identify potentially relevant studies that our initial search may have missed.[16]

We included all estimates of sibling correlations in educational attainment, subject to the following conditions. We excluded:

- Twin samples. Twins differ systematically from other siblings (Smits and

[15]This search nonetheless yielded some studies not written in English, and of these we reviewed the studies in languages spoken by the research team (Dutch, English, French, German, Latvian, Russian, and Uzbek).

[16]We searched Google Scholar, the Education Resources Information Center, and EconLit, using a less comprehensive set of terms. We conducted a backward search, following up references in relevant studies and existing narrative reviews of related topics (Björklund and Salvanes, 2011; Black and Devereux, 2011; Griliches, 1979). We conducted a forward search

Monden, 2011).[17]

- Samples in which a non-negligible proportion of children had not yet completed their education.[18]

- Samples limited to children living in the parental home. Household formation or migration (and thus missingness) are likely to correlate with education.

- Estimates based on educational achievement (e.g. test scores), cognitive ability, enrolment, or other related constructs but not actual *attainment*.

- Categorical measures of attainment that aggregate substantively distinct levels.[19]

- Categorical measures of attainment for which estimates are calculated as intraclass correlation coefficients. 'ICCs are often compared between groups to show the degree to which sibling differences vary between groups: we show that when the outcome is categorical these comparisons are invalid.' (Breen and Ermisch, 2021, p. 497).

- Estimates based on pooled samples from multiple countries.

- Estimates specific to any subgroup other than brothers or sisters. For example we excluded estimates:

  - based solely on brother–sister pairs;
  - disaggregated by sibship size;
  - only reported separately across subgroups such as short birth-spacing/long birth-spacing, or respondents and their older brothers/respondents

by using Google Scholar to identify studies citing studies we had already included. We also contacted selected authors in the field to obtain any unpublished studies.

[17]For a meta-analysis of twin studies on educational attainment see Branigan *et al.* (2013).

[18]Whether this is explicitly stated in or deducible from the study, or would be expected based on the age of the sample.

[19]This means we exclude, for instance, binary measures of tertiary attainment. Almost all of our estimates measure attainment in years or year-equivalents. We include six estimates

and their younger brothers.[20]

- Estimates for overlapping birth cohorts within the same study. For example if a study presented estimates for three birth cohorts 1980–1989, 1985–1994, and 1990–1999, we excluded the middle one.

- Estimates that aggregated over multiple birth cohorts, *where disaggregated estimates were also available in the same study.*

- Estimates from models including any covariates beyond sex, age, birth year, and birth order. In practice it was rare for any covariates to be included.

- Studies exactly restating estimates originally reported elsewhere in the literature.

- Earlier versions of studies of which we found published or updated versions.

To be clear, note that this implies that we *included*:

- Separate *brother*, *sister*, and *all siblings* (i.e. non gender-specific) estimates for the same sample.

- Estimates from different studies using the same sample (as long as one is not simply a restatement of another, such as in a handbook chapter or review).

- The most disaggregated estimates in terms of birth cohort, among those presented in a study.

C List of included studies

Adermon, Adrian, 2013. 'Sibling Spillover in Education. Causal Estimates from a Natural Experiment.' In *Essays on the Transmission of Human Capital and the Impact of Technological Change*, pp. 31–62. Uppsala: Department of Economics, Uppsala University.

based on categorical measures, which all distinguish at least seven levels.

[20]We made exceptions for two cases: Duncan *et al.* (1972) only present estimates disaggregated by race, and we included estimates for both Black and White samples, recording them as de facto different datasets; Wells (1995) only presents estimates separately for one- and two-parent families and we included the latter, which represented over 80% of the sample. None of these estimates are nationally representative, so they are excluded from our meta-regression analyses.

Ashenfelter, Orley and David J. Zimmerman, 1997. ‘Estimates of the Returns to Schooling from Sibling Data. Fathers, Sons, and Brothers.’ *Review of Economics and Statistics* 79(1): 1–9. doi: 10.1162/003465397556421.

Beenstock, Michael, 2008. ‘Deconstructing the Sibling Correlation. How Families Increase Inequality.’ *Journal of Family and Economic Issues* 29(3): 325–345. doi: 10.1007/s10834-008-9114-y.

Benin, Mary Holland and David R. Johnson, 1984. ‘Sibling Similarities in Educational Attainment. A Comparison of Like-Sex and Cross-Sex Sibling Pairs.’ *Sociology of Education* 57(1): 11–21. doi: 10.2307/2112464.

Bingley, Paul and Lorenzo Cappellari, 2019. ‘Correlation of Brothers’ Earnings and Intergenerational Transmission.’ *Review of Economics and Statistics* 101(2): 370–383. doi: 10.1162/rest_a_00753.

Bingley, Paul, Lorenzo Cappellari, and Konstantinos Tatsiramos, 2021. ‘Family, Community, and Long-Term Socio-Economic Inequality. Evidence from Siblings and Youth Peers.’ *Economic Journal* 131(636): 1515–1554. doi: 10.1093/ej/ueaa121.

Björklund, Anders and Markus Jäntti, 2012. ‘How Important Is Family Background for Labor-Economic Outcomes?’ *Labor Economics* 19(4): 465–474. doi: 10.1016/j.labeco.2012.05.016.

Björklund, Anders, Markus Jäntti, and Matthew J. Lindquist, 2009. ‘Family Background and Income During the Rise of the Welfare State. Brother Correlations in Income for Swedish Men Born 1932–68.’ *Journal of Public Economics* 93(5-6): 671–680. doi: 10.1016/j.jpubeco.2009.02.006.

Björklund, Anders and Kjell G. Salvanes, 2011. ‘Education and Family Background. Mechanisms and Policies.’ In Hanushek, Eric A., Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, volume 3, pp. 201–247. Amsterdam: Elsevier. doi: 10.1016/B978-0-444-53429-3.00003-X.

Bound, John, Zvi Griliches, and Bronwyn H. Hall, 1986. ‘Wages, Schooling, and IQ of Brothers and Sisters. Do the Family Factors Differ?’ *International Economic Review* 27(1): 77–105. doi: 10.2307/2526608.

Bredtmann, Julia and Nina Smith, 2018. ‘Inequalities in Educational Outcomes. How Important Is the Family?’ *Oxford Bulletin of Economics and Statistics* 80(6): 1117–1144. doi: 10.1111/obes.12258.

Brittain, John A., 1977. *The Inheritance of Economic Status.* Washington, DC: Brookings Institution.

Bronars, Stephen G. and Gerald S. Oettinger, 2006. 'Estimates of the Return to Schooling and Ability. Evidence from Sibling Data.' *Labor Economics* 13(1): 19–34. doi: 10.1016/j.labeco.2004.07.003.

Cawley, John, Euna Han, Jiyoon Kim, and Edward C. Norton, 2020. 'Sibling Correlation in Educational Attainment. A Test of Genetic Nurture.' *NBER Working Paper* 27336. doi: 10.3386/w27336.

Chamberlain, Gary and Zvi Griliches, 1975. 'Unobservables with a Variance Components Structure. Ability, Schooling, and the Economic Success of Brothers.' *International Economic Review* 16(2): 422–49. doi: 10.2307/2525824.

Conley, Dalton and Rebecca Glauber, 2007. 'Family Background, Race, and Labor Market Inequality.' *Annals of the American Academy of Political and Social Science* 609(1): 134–152. doi: 10.1177/0002716206296090.

——, 2008. 'All in the Family? Family Composition, Resources, and Sibling Similarity in Socioeconomic Status.' *Research in Social Stratification and Mobility* 26(4): 297–306. doi: 10.1016/j.rssm.2008.08.003.

Corcoran, Mary and Linda P. Datcher, 1981. 'Intergenerational Status Transmission and the Process of Individual Attainment.' In Hill, Martha S., Daniel H. Hill, and James N. Morgan, eds., *Five Thousand American Families. Patterns of Economic Progress*, volume IX, pp. 169–206. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Corcoran, Mary, Christopher Jencks, and Michael Olneck, 1976. 'The Effects of Family Background on Earnings.' *American Economic Review* 66(2): 430–435. doi: 10.2307/1817256.

de Graaf, Paul M., 1986. 'The Impact of Financial and Cultural Resources on Educational Attainment in the Netherlands.' *Sociology of Education* 59(4): 237–246. doi: 10.2307/2112350.

de Graaf, Paul M. and Johannes J. Huinink, 1992. 'Trends in Measured and Unmeasured Effects of Family Background on Educational Attainment and Occupational Status in the Federal Republic of Germany.' *Social Science Research* 21(1): 84–112. doi: 10.1016/0049-089X(92)90019-D.

Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan, 1972. *Socioeconomic Background and Achievement.* New York: Seminar Press.

Ermisch, John and Chiara Pronzanto, 2011. 'Causal Effects of Parents' Education on Children's Education.' In Smeeding, Timothy M., Robert Erikson, and Markus Jäntti, eds., *Persistence, Privilege, and Parenting. The Comparative Study of Intergenerational Mobility*, pp. 237–260. New York: Russell Sage.

Erola, Jani, 2012. *The Life Course Variation of Sibling Correlations According to Class and Education.* University of Turku. doi: 10.2139/ssrn.2133753.

Grätz, Michael, Kieron J. Barclay, Øyvind Wiborg, Torkild H. Lyngstad, Aleksi Karhula, Jani Erola, Patrick Präg, Thomas Laidley, and Dalton Conley, 2021. 'Sibling Similarity in Education Across and Within Societies.' *Demography* 58(3): 1011–1037. doi: 10.1215/00703370-916402.

Hauser, Robert M., 1984. 'Some Cross-Population Comparisons of Family Bias in the Effects of Schooling on Occupational Status.' *Social Science Research* 13(2): 159–187. doi: 10.1016/0049-089X(84)90019-X.

Hauser, Robert M. and David L. Featherman, 1976. 'Equality of Schooling. Trends and Prospects.' *Sociology of Education* 49(2): 99–120. doi: 10.2307/2112516.

Hauser, Robert M. and William H. Sewell, 1986. 'Family Effects in Simple Models of Education, Occupational Status, and Earnings. Findings from the Wisconsin and Kalamazoo Studies.' *Journal of Labor Economics* 4(3, Part 2): S83–S115. doi: 10.1086/298122.

Hauser, Robert M., Jennifer T. Sheridan, and John Robert Warren, 1999. 'Socioeconomic Achievements of Siblings in the Life Course. New Findings from the Wisconsin Longitudinal Study.' *Research on Aging* 21(2): 338–378. doi: 10.1177/0164027599212008.

Hällsten, Martin and Max Thaning, 2022. 'Wealth as One of the "Big Four" SES Dimensions in Intergenerational Transmissions.' *Social Forces* 100(4): 1533–1560. doi: 10.1093/sf/soab080.

Jæger, Mads Meier, 2012. 'The Extended Family and Children's Educational Success.' *American Sociological Review* 77(6): 903–922. doi: 10.1177/0003122412464040.

Lecavelier des Etangs-Levallois, Céline and Arnaud Lefranc, 2017. 'Sibling Correlations in Terms of Education, Profession, and Earnings in France.' *THEMA Working Paper* 2017-12.

Lee, Kristen Schultz, 2009. 'Competition for Resources. A Reexamination of Sibship Composition Models of Parental Investment.' *Journal of Marriage and Family* 71(2): 263–277. doi: 10.1111/j.1741-3737.2009.00598.x.

Levine, David I. and Bhashkar Mazumder, 2007. 'The Growing Importance of Family. Evidence from Brothers' Earnings.' *Industrial Relations* 46(1): 7–21. doi: 10.1111/j.1468-232X.2007.00455.x.

Lindahl, Lena, 2011. 'A Comparison of Family and Neighborhood Effects on Grades, Test Scores, Educational Attainment, and Income. Evidence from Sweden.' *Journal of Economic Inequality* 9(2): 207–226. doi: 10.1007/s10888-010-9144-1.

Marks, Gary N. and Irma Mooi-Reci, 2016. 'The Declining Influence of Family Background on Educational Attainment in Australia. The Role of Measured and Unmeasured Influences.' *Social Science Research* 55: 171–185. doi: 10.1016/j.ssresearch.2015.10.002.

Mazumder, Bhashkar, 2008. 'Sibling Similarities and Economic Inequality in the US.' *Journal of Population Economics* 21: 685–701. doi: 10.1007/s00148-006-0127-2.

——, 2011. 'Family and Community Influences on Health and Socioeconomic Status. Sibling Correlations Over the Life Course.' *BE Journal of Economic Analysis and Policy* 11(3): 1–21. doi: 10.2202/1935-1682.2876.

Müller, Walter, 1972. 'Family Background, Education, and Career Mobility.' *Social Science Information* 11(5): 223–255. doi: 10.1177/053901847201100510.

Nybom, Martin and Jan Stuhler, 2019. 'Steady-State Assumptions in Intergenerational Mobility Research.' *Journal of Economic Inequality* 17(1): 77–97. doi: 10.1007/s10888-019-09412-y.

Olneck, Michael R., 1977. 'On the Use of Sibling Data to Estimate the Effects of Family Background, Cognitive Skills, and Schooling. Results from the Kalamazoo Brothers Study.' In Taubman, Paul, ed., *Kinometrics. Determinants of Socioeconomic Success Within and Between Families*, pp. 125–162. Amsterdam: North-Holland.

Pfeffer, Fabian T., Alexandra Killewald, and Andreja Siliunas, 2016. *The Concentration of Wealth within Family Lineages and Intergenerational Transfers*. PSID Conference 'New Directions in the Study of Intergenerational Transfers and Time Use in Later Life'.

Raaum, Oddbjørn, Kjell G. Salvanes, and Erik Sørensen, 2006. 'The Neighborhood Is Not What It Used to Be.' *Economic Journal* 116(508): 200–222. doi: 10.1111/j.1468-0297.2006.01053.x.

Sacerdote, Bruce, 2007. 'How Large Are the Effects From Changes in Family Environment? A Study of Korean American Adoptees.' *Quarterly Journal of Economics* 122(1): 119–157. doi: 10.1162/qjec.122.1.119.

Scarr, Sandra and Richard A. Weinberg, 1994. ‘Educational and Occupational Achievements of Brothers and Sisters in Adoptive and Biologically Related Families.’ *Behavior Genetics* 24(4): 301–325. doi: 10.1007/BF01067532.

Schnitzlein, Daniel 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. doi: 10.1007/s00148-013-0468-6.

Sewell, William H. and Robert M. Hauser, 1977. ‘On the Effects of Family and Family Structure on Achievement.’ In Taubman, Paul, ed., *Kinometrics. Determinants of Socioeconomic Success Within and Between Families*, pp. 256–286. Amsterdam: North-Holland.

Sieben, Inge, 2001. *Sibling Similarities and Social Stratification. The Impact of Family Background across Countries and Cohorts.* Nijmegen: ICS/Radboud University.

Sieben, Inge and Paul M. de Graaf, 2001. ‘Testing the Modernization Hypothesis and the Socialist Ideology Hypothesis. A Comparative Sibling Analysis of Educational Attainment and Occupational Status.’ *British Journal of Sociology* 52(3): 441–467. doi: 10.1080/00071310120071133.

——, 2003. ‘The Total Impact of the Family on Educational Attainment. A Comparative Sibling Analysis.’ *European Societies* 5(1): 33–68. doi: 10.1080/1461669032000057668a.

——, 2004. ‘Schooling or Social Origin? The Bias in the Effect of Educational Attainment on Social Orientations.’ *European Sociological Review* 20(2): 107–122. doi: 10.1093/esr/jch011.

Sieben, Inge, Johannes Huinink, and Paul M. de Graaf, 2001. ‘Family Background and Sibling Resemblance in Educational Attainment. Trends in the Former FRG, the Former GDR, and the Netherlands.’ *European Sociological Review* 17(4): 401–430. doi: 10.1093/esr/17.4.401.

Sirniö, Outi, Hannu Lehti, Michael Grätz, Kieron Barclay, and Jani Erola, 2020. ‘The Pattern of Educational Inequality. The Contribution of Family Background on Levels of Education Over Time and Across Four Countries.’ doi: 10.31235/osf.io/nupfs.

Solon, Gary, Marianne E. Page, and Greg J. Duncan, 2000. ‘Correlations between Neighboring Children in Their Subsequent Educational Attainment.’ *Review of Economics and Statistics* 82(3): 383–392. doi: 10.1162/003465300558885.

Sweetser, Dorrian Apple, 1973. *Urban Norwegians. Kinship Networks, and Sibling Mobility.* Oslo: Institute of Applied Social Research.

Sweetser, Dorrian Apple and Patrick McDonnell, 1978. 'Social Origins, Education, and Fraternal Mobility.' *American Journal of Sociology* 83(4): 975–982. doi: 10.1086/226640.

Thaning, Max and Martin Hällsten, 2020. 'The End of Dominance? Evaluating Measures of Socio-Economic Background in Stratification Research.' *European Sociological Review* 36(4): 533–547. doi: 10.1093/esr/jcaa009.

Toka, Gabor and Jaap Dronkers, 1996. 'Sibling Resemblance in Educational Attainment, Occupational Prestige, and Wealth in Hungary during the Communist Regime.' *European Sociological Review* 12(3): 251–269. doi: 10.1093/oxfordjournals.esr.a018191.

Torvik, Fartein Ask, Espen Moen Eilertsen, Laurie J. Hannigan, Rosa Cheesman, Laurence J. Howe, Per Magnus, Ted Reichborn-Kjennerud, Ole A. Andreassen, Pål R. Njølstad, Alexandra Havdahl, and Eivind Ystrom, 2022. 'Modeling Assortative Mating and Genetic Similarities Between Partners, Siblings, and In-Laws.' *Nature Communications* 13(1): 1108. doi: 10.1038/s41467-022-28774-y.

van Eijck, Koen, 1996. *Family and Opportunity. A Sibling Analysis of the Impact of Family Background on Education, Occupation, and Consumption.* Tilburg: Tilburg University Press.

Wells, Thomas, 1995. 'Does Family Background Affect Educational Attainment Differently According to Family Structure, Birth Order, and Sex?' *NSFH Working Paper* 70.

Wiborg, Øyvind N. and Marianne N. Hansen, 2018. 'The Scandinavian Model During Increasing Inequality. Recent Trends in Educational Attainment, Earnings, and Wealth among Norwegian Siblings.' *Research in Social Stratification and Mobility* 56: 53–63. doi: 10.1016/j.rssm.2018.06.006.

Wong, Maisy, 2019. 'Intergenerational Mobility in Slums. Evidence from a Field Survey in Jakarta.' *Asian Development Review* 36(1): 1–19. doi: 10.1162/adev_a_00121.

D Data extraction notes

This section is intended to clarify the choices we made when confronted with difficulties or ambiguities in the data extraction process. It is not exhaustive, and mainly focuses on our decisions in cases where we found an apparently includable sibling correlation estimate, but were unable to straightforwardly find details such as sample size or the birth years of the sample. Except for minor details, we attempted to contact the author(s) for clarification in these cases. We note cases in which information was given to us by authors.

In some instances our attempts to contact the author(s) for clarification were unsuccessful. Therefore, as we also record here, some estimates were coded as 'Model unclear', and in some cases we imputed the sample size. These imputations were based on information such as sample sizes reported in other sibling correlation studies using the same dataset, our attempts to replicate the published correlations using the original data, and (because our sample sizes were the number of families) the number of individuals (i.e. siblings) in the sample, which usually was reported if the number of families was not. Similarly we sometimes had to make a reasonable estimate of other details such as the span of sample birth years. In all cases we were able to estimate imputed sample size with a high degree of confidence; it is also worth noting that given our use of random-effects models, only extremely large errors could be expected to noticeably change our results.

East Germany, West Germany, Germany The classification of German samples is complicated by history. Different individuals grew up and completed their education in the former East Germany, the former West Germany, and unified Germany; some were partway through their educational careers when unification occurred. Further, some samples are of adults in (unified) Germany and therefore represent a mix of people who grew up under different systems.

In our dataset of estimates we distinguish a) East Germany (estimates from the East German Surveys of the German Life History Study (Mayer, 2015)), b) West Germany (estimates from the West German Surveys of the German Life History Study and the Konstanz sample from Müller (1972)), and c) Germany (estimates from the German Socio-Economic Panel). a) and b) refer exclusively to cohorts that grew up in the former East and West Germany respectively. Note that c) includes both cohorts that grew up wholly or partially in a unified Germany and cohorts that comprise a mix of individuals who grew up in the East and West (more in the West, by about four to one) and now represent the middle-aged and older population of Germany.

In our meta-regression analyses we combine b) and c) and categorize these estimates as referring to Germany. We thereby contrast the particularly distinctive East German experience with the other types of sample mentioned.

The *dtfl* variable in our dataset of estimates A value of 1 on the variable *dtfl* (dataset flag) indicates that the dataset a sample comes from overlaps with the dataset of another estimate (i.e. includes some of the same individuals), *and* specifically, unlike in most cases, that this is not captured by their having the same value of the variable dataset. To illustrate: two samples (with identical or overlapping birth years) from e.g. the PSID would normally be identified as containing at least some of the same individuals by virtue of (the variables indicating the cohort birth year span and) the variable dataset. This would straightforwardly be 'Panel Study of Income Dynamics' in both cases. However, some estimates are for samples drawn from a combination of datasets. The dataset variable may therefore fail to identify overlap between samples in such cases. For example: 'Social Mobility Study Hungary 1983' and 'Social Mobility Study Hungary 1983 and Hungarian Social Mobility and Life History Survey 1992' are different. *dtfl* flags estimates where the sample overlaps with the sample for another estimate, but this overlap cannot be identified based on identical values of dataset.

dtfl is useful because it identifies where, in writing our data preparation code, we cannot simply rely on the variable dataset to distinguish separate datasets, and instead have to deal with the estimates case-by-case.

Most $dtfl = 1$ cases come from studies by Sieben (2001); Sieben and de Graaf (2003) because these authors produce many estimates from pooled datasets. Another instance is that the International Social Science Survey Australia 1984-2001 Pooled File used by Marks and Mooi-Reci (2016) contains within it the Australian National Social Science Survey 1989/1990 used by Sieben and de Graaf (2003). $dtfl = 1$ cases arise with respect to four countries: Australia, Hungary, the Netherlands, and the US.

Adermon (2013) The author kindly provided access to the full text of this study, which is chapter 2 of his doctoral dissertation.

Björklund and Salvanes (2011) Model unclear—no information given about how the estimates were arrived at. Sample sizes imputed.

Björklund *et al.* (2009) Sample sizes imputed.

Bredtmann and Smith (2018) Sample sizes for the brother and sister correlations from register data are unclear. For these two estimates, we impute the number of families as half the (reported) number of families for the all siblings correlation. Since singletons are included in the samples, this represents a conservative estimate.

Cawley *et al.* (2020) has in the meantime been published as Cawley *et al.* (2023).

Chamberlain and Griliches (1975) The estimate from this study is reported in Table 1 of Griliches (1979).

Conley and Glauber (2007) Beginning of sample birth year range unclear. 'We select adult respondents aged twenty-five and older who were head of their household or married to the head of household in any (or all) years between 1983 and 2001.' Table 1 gives mean age as ca. 36 with SD ca. 7, suggesting a plausible maximum age of ca. 50; this also fits with the presentation of results in later tables separately for those aged 30 and under, 31–40, and over 40. So we have coded *coh_start* as $1983 - -50 = 1933$.

Conley and Glauber (2008) Sample description is almost identical to Conley and Glauber (2007), so we have coded *coh_start* as 1933 for this study too.

Grätz *et al.* (2021) The German sample (from the SOEP) were born 1976–1989. We have recorded the country as 'Germany' rather than 'East Germany' or 'West Germany,' since most of this cohort will have undertaken most of their education in a unified Germany.

Hällsten and Thaning (2022) The authors kindly provided us with sample size in terms of numbers of families.

Hauser and Featherman (1976) Sample sizes imputed.

Hauser and Sewell (1986) Unclear whether the sibling correlations are based on the 'pooled sample of complete and incomplete data' ($532+928 = 1{,}460$ brother pairs) or the 'complete' sample ('532 pairs meet our criteria for inclusion in the analysis; hereafter, we refer to this as the complete sample'). We conservatively assume $N = 532$.

Jæger (2012) Sample size for the NLSY79-CYA imputed. Both the number of individuals and the number of families in the WLS sample are stated. Only the number of individuals in the NLSY79-CYA sample is stated. We impute number of families by assuming the same average family size in the two samples.

Sample birth year ranges unclear.

- WLS sample: beginning of range (*coh_start*) unclear. Table 1 gives age in 2004 as mean 38.2, SD 5.4. We have taken mean + 2SD as an approximation of the maximum age. This yields 1955 for *coh_start*. (Note this also makes sense in light of the fact that these are the children of WLS respondents, 'born in or around 1939', and of their siblings.)
- NLSY79-CYA sample: 'The NLSY-CYA includes all children born to women who participated in the National Longitudinal Survey of Youth 1979.' Sparsely described. Citations are to 2006 editions of user guides; data collected every two years from 1986–2018, so seems fair to assume

they are using 2006 data. Given sample restriction to those aged at least 25, we estimate *coh_end* is $2006 - 25 = 1981$. NLSY79 respondents were 14–21 at end of 1978, therefore born 1957–1964. NLS website says: 'Age of the NLSY79 Child & Young Adult cohorts: Born between 1970 and 2014. At the time of the first interview in 1986, child ages ranged from 0–23 years [sic].' For *coh_start*, we take the website's most direct statement and say 1970.

Lecavelier des Etangs-Levallois and Lefranc (2017) Sample sizes for the estimates (from Table 8) are not given in the manuscript, but they are reported in Table 9 of an earlier version (Lecavelier des Etangs-Levallois and Lefranc, 2015).

Levine and Mazumder (2007) Sample size unclear from the article itself: on the one hand, the text (p. 18) makes clear that the sample for the NLS66 estimate is 341 pairs of brothers (but does not state a sample size for the NLSY79 estimate), while on the other hand, the note to Table 5 says that the samples used are the same as reported in Table 3—where, however, education is not one of the outcomes (and the sample sizes given are much larger). An earlier version of the article resolves the ambiguity (Mazumder and Levine, 2004). Table 6 in the working paper version reports the same estimates with sample sizes—including 341 for the NLS66 sample.

Mazumder (2008) Specifically for the education outcome, numbers of families are not stated. We impute the numbers of families given in descriptions of samples for other outcomes (the corresponding number of individuals in each case is very similar to the education sample's number of individuals, which is reported).

Mazumder (2011) Sample size imputed. Only the number of individuals underlying the specific estimates is reported. We impute number of families based on the family size implied by the statement 'The full adult sample includes 3,265 individuals from 1,355 families' (p. 6).

Müller (1972) Sample size ($N = 200$) imputed. The number of families in the sample as a whole is given as 398 (p. 224) (assuming each male 'born in 1936 or in the first three months of 1937' comes from a different family); but the ICC is calculated for 'all families of the sample having at least two sons' (p. 241).

Nybom and Stuhler (2019) The authors kindly provided us with the sample size in terms of number of families. They also clarified that singletons were excluded and that they calculated a Pearson correlation weighting each family equally regardless of the number of siblings.

Raaum *et al.* (2006) The authors kindly provided us with sample sizes in terms of numbers of families.

Sacerdote (2007) We include the estimate for biological siblings and exclude the estimate referring to adoptive siblings.

Schnitzlein (2014) This German sample is from the SOEP and appears to span the birth years 1953–1983 (see below). Although it therefore includes a minority who grew up and completed most of their education in the former East Germany, we record the country as 'Germany.'

Sample birth year range not totally clear. The study uses SOEP 2002–2008; minimum age is 25 (both stated under Table 6; see also footnote 27). Table 2 (descriptives) says 'Individuals are between 31 and 49 years of age' with mean ca. 35–38. So *coh_end* is $2008 - 25 = 1983$; for *coh_start* it seems reasonable to say $2002 - 49 = 1953$ based on the Table 2 note.

Sieben and de Graaf (2001) Unlike most other studies (including by the same authors), the cohorts given in the text are labor market cohorts: 'we constructed labor market cohorts with a range of 15 years (1916–1930, 1931–1945, 1946–1960, 1961–1975, and 1976–1990). These labor market cohorts are based on the year in which the brothers made the transition into the labor market ... The starting year in the occupational career is defined as the birth year added by the school leaving age plus two.' No further information is given within the study about school leaving ages or birth years/birth cohorts. We have approximated birth year by assuming a school-leaving age of 13 across all countries and periods, i.e. we have assumed birth year is the given labour market cohort year minus 15.

Sieben *et al.* (2001) Note that we have taken the 'Duration' correlations, as these are based on transforming education into year-equivalents.

Sieben and de Graaf (2003) We used the WebPlotDigitizer tool to extract sibling correlations from Figure 1. We identified country–cohorts by matching values of per capita energy use from the figure to those listed in Appendix B.

Note that unlike most other studies, many estimates combine data from multiple datasets (see their Appendix A). For this reason our dataset of estimates contains two variables describing the dataset(s) on which each estimate is based: *dataset_infull*, and *dataset*. The two are identical except in the case of estimates from Sieben and de Graaf (2003) for Czechoslovakia and Hungary.

dataset abbreviates some very long strings, so that it is slightly easier to work with. For example:

dataset: SMS HUN83; TARKI-I86; HUN SMLHS92; SMOC HUN73; SSEE GPS93 HUN

dataset_infull: Social Mobility Study Hungary 1983; TARKI-I 1986; Hungarian Social Mobility and Life History Survey 1992; Social Mobility and Occupational Changes in Hungary 1973; Social Stratification in Eastern Europe after 1989: General Population Survey 1993 Hungary

Sirniö *et al.* (2020) Model unclear—unclear whether singletons were included or not.

Sample sizes imputed.

Estimates are reported for overlapping cohorts: sibling correlations are given for 31–39-year-olds at (mostly) five-year intervals (e.g. 1990, 1995, 2000, 2005). So that individuals were not all included in samples for two different estimates, we included every other estimate, i.e. estimates at ten-year intervals. We (arbitrarily) kept estimates from 1990, 2000, and 2010, since this more neatly fits the way we assign estimates to decadal birth cohorts (e.g. the 1990 sample are born 1951–1959 and may be coded as a 1950s cohort, whereas the 1995 sample are born 1956–1964).

The sample for Germany is from the SOEP and includes cohorts born in the 1960s and 1970s. Although this therefore includes a minority who grew up and completed most of their education in the former East Germany, we record the country as 'Germany.'

Solon *et al.* (2000) This study presents four estimates differentiated only by the weighting scheme used. We included the estimate which gives equal weight to each family regardless of sibship size.

Sweetser (1973) Sample birth year range unclear. '[S]ample of 500 urban Norwegians, over 19 years of age' from a '1967 master sample'; but 'interviews were carried out [in] 1970'. So *coh_end* is $1970 - 20 = 1950$. Table 1-1 on p. 5 breaks the sample down by age categories with the oldest being '70 and over'. We take 80 as a plausible maximum age. So *coh_start* is $1970 - 80 = 1890$.

Thaning and Hällsten (2020) The authors kindly provided us with sample sizes in terms of numbers of families.

Toka and Dronkers (1996) A sample size of 831 families is given for the sample as a whole. However this sample is disaggregated into five cohorts and correlations are reported only for the oldest three. Because the cohorts are of similar length (i.e. range of birth years) and no further information is given about the distribution of the overall sample between these cohorts, we impute sample size for each of these three cohorts as 831/5.

Torvik *et al.* (2022) Sample birth year range unclear. We estimate 1959–1978 based on the following: education is measured at age 30, and age is mean ca. 30 with SD ca. 5 for both men and women. Data collection appears to have been continuous over 1999–2008.

Wiborg and Hansen (2018) The authors kindly provided us with sample sizes in terms of numbers of families.

Wong (2019) Age range (and thus sample birth year range) unclear. Survey conducted in 2016, minimum age 18, maximum unclear but 50th percentile is 49, 75th percentile is 58. We impute 65 for maximum age and hence record birth year range as 1951–1998.

E Publication bias

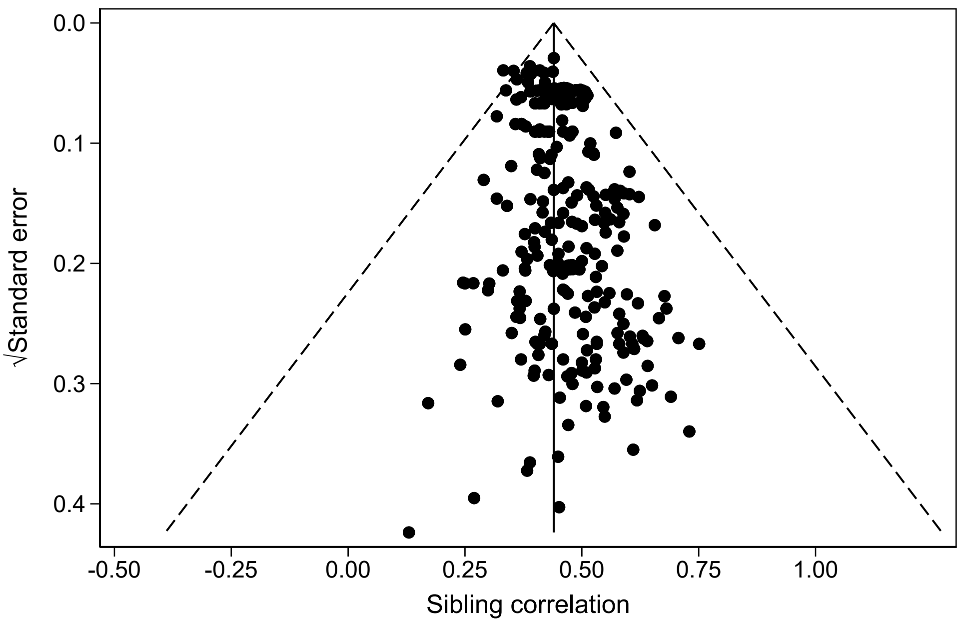

Figure S2: Funnel plot of all sibling correlation estimates meeting inclusion criteria ($N = 300$) against the square root of the standard error, as a measure of study precision.

The funnel plot in Figure S2 is approximately symmetrical in appearance, which gives no indication of likely publication bias. This is in line with our expectation that such bias is unlikely in this context, which we hold for two main reasons.

First, sibling correlations are invariably statistically significant—or to be more specific, across our literature search we have come across none that were not statistically significant. Second, sibling correlations are often not the focus or main result of a study. Rather one usually finds them reported as descriptive or supplementary information. Thus it seems unlikely that researchers would fail to publish—or fear failing to publish and therefore consciously or unconsciously manipulate—a study because of an unexpectedly low or non-significant sibling correlation.

These considerations illustrate how our case differs from more common scenarios in meta-analysis in which publication bias is a major concern. That is, where the estimate of interest is the outcome of, for instance, a randomised controlled trial, and researchers and publishers may have incentives to not publish results that are non-significant, or are of an unexpected or unusual magnitude.

F Further overview of the sibling correlations literature

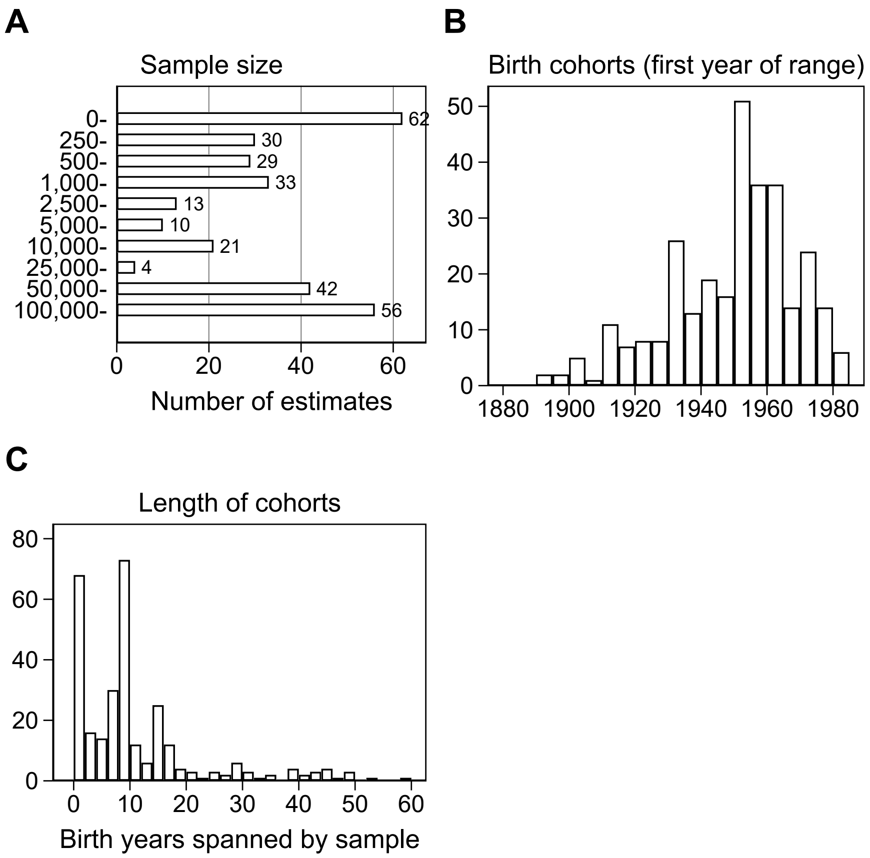

Figure S3: An overview of the sibling correlations literature (further detail). *Panel A:* distribution of sample sizes. *Panel B:* distribution of birth cohorts, as indicated by first year of birth among the sample. *Panel C:* distribution of cohort length, meaning number of birth years spanned by individuals in the sample.

Note to Panel A: sample size refers to the number of families in each case. Note to Panel B: for clarity, this excludes an outlier: the Gorseline Brothers sample (US; individuals born 1855–1908), for which an estimate of .24 based on 156 families is reported in Chamberlain and Griliches (1975).

G Forest plot of single estimated sibling correlations per sample, by country

Figure S4 shows predicted values based on a meta-regression of 253 sibling correlation estimates from the literature on sample, sibship type, and model type. The forest plot shows the predicted correlation for each sample (i.e. dataset–

cohort combination), holding sibship type at *All siblings* and model type at *Pearson's r*.

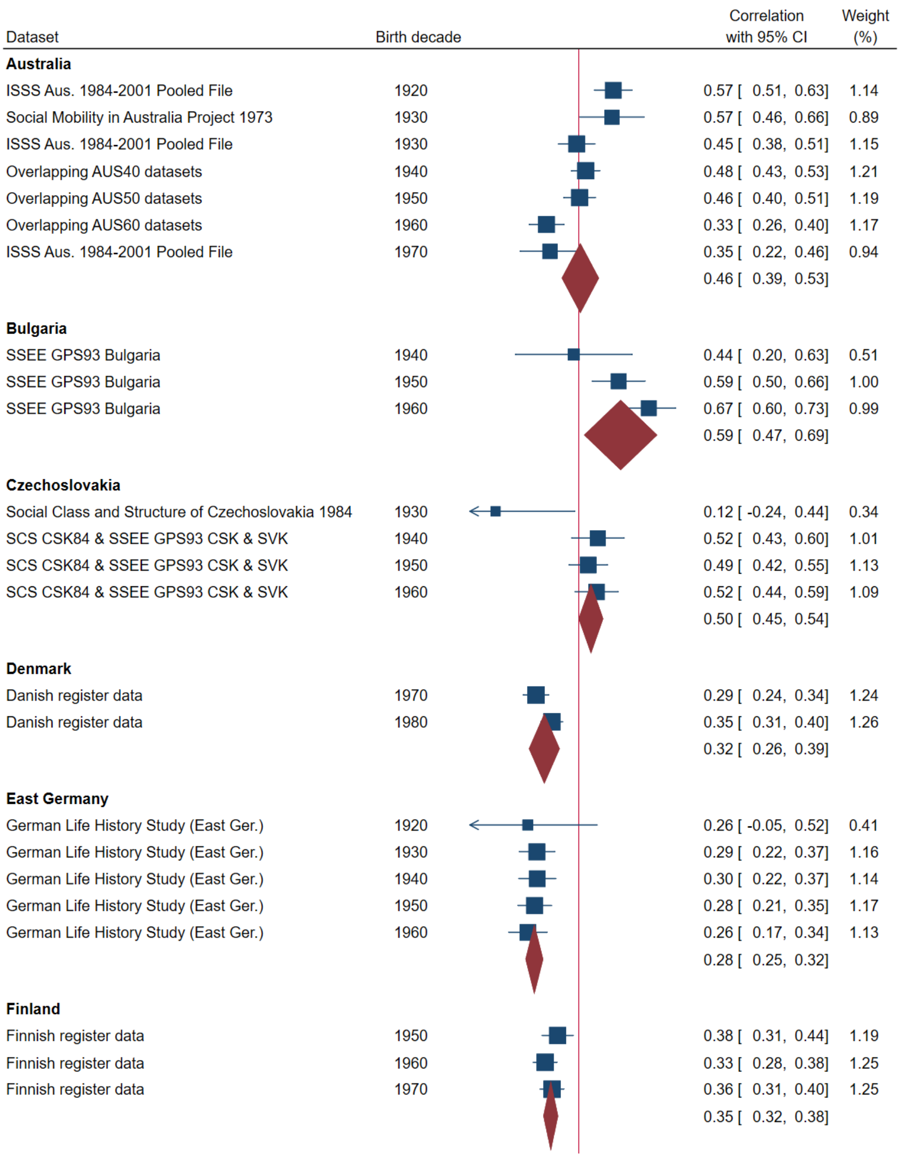

Figure S4: Forest plot of single estimated sibling correlations per sample, by country *(continued overleaf)*

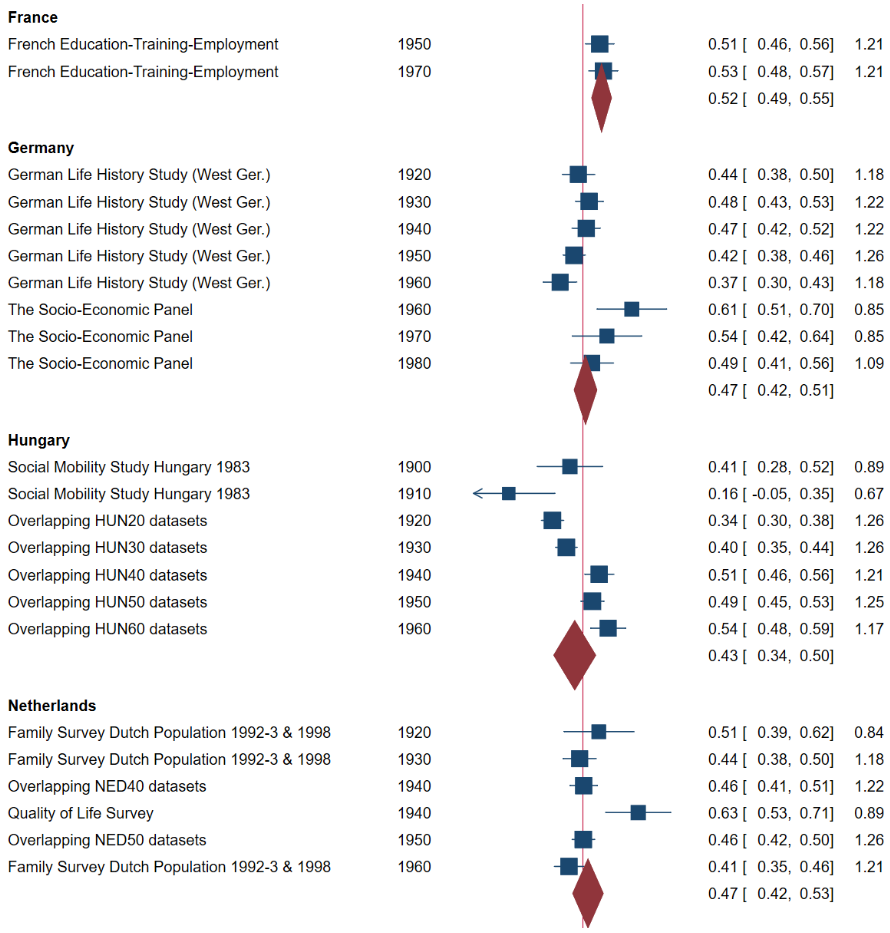

Figure S4: Forest plot of single estimated sibling correlations per sample, by country *(continued)*

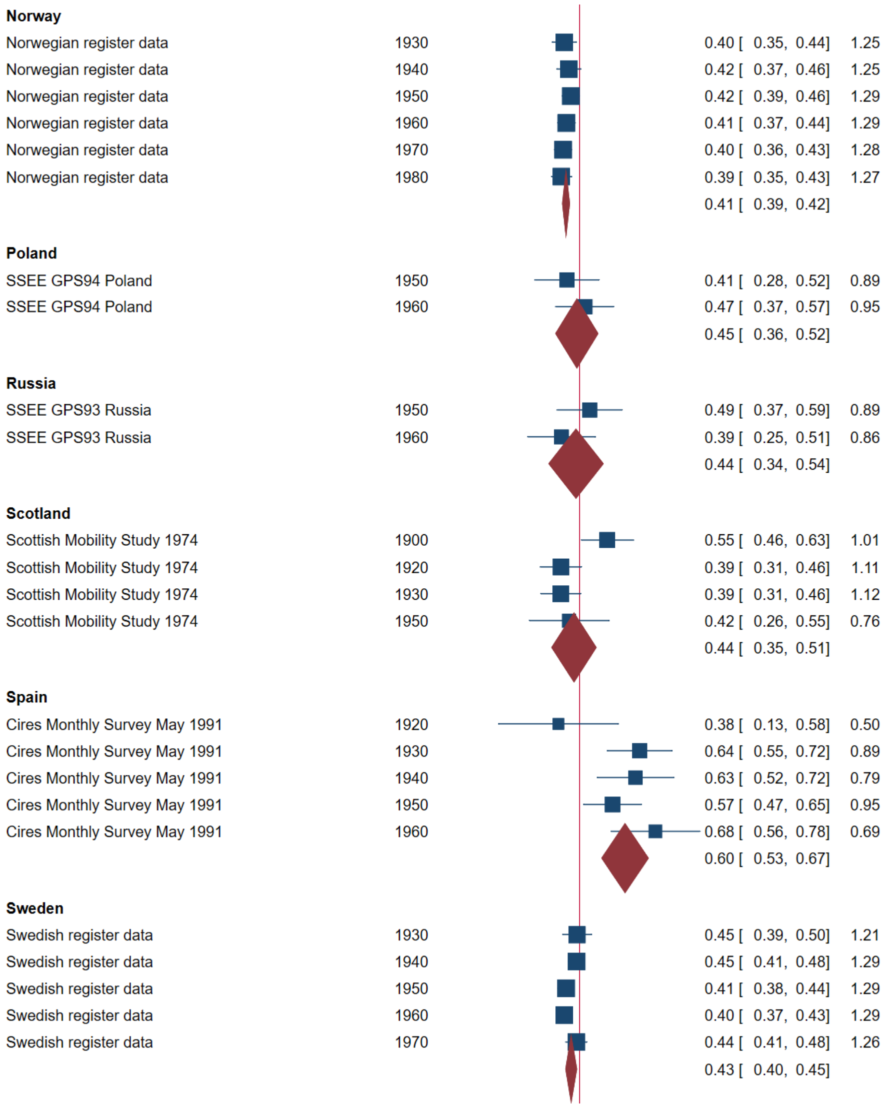

Figure S4: Forest plot of single estimated sibling correlations per sample, by country *(continued)*

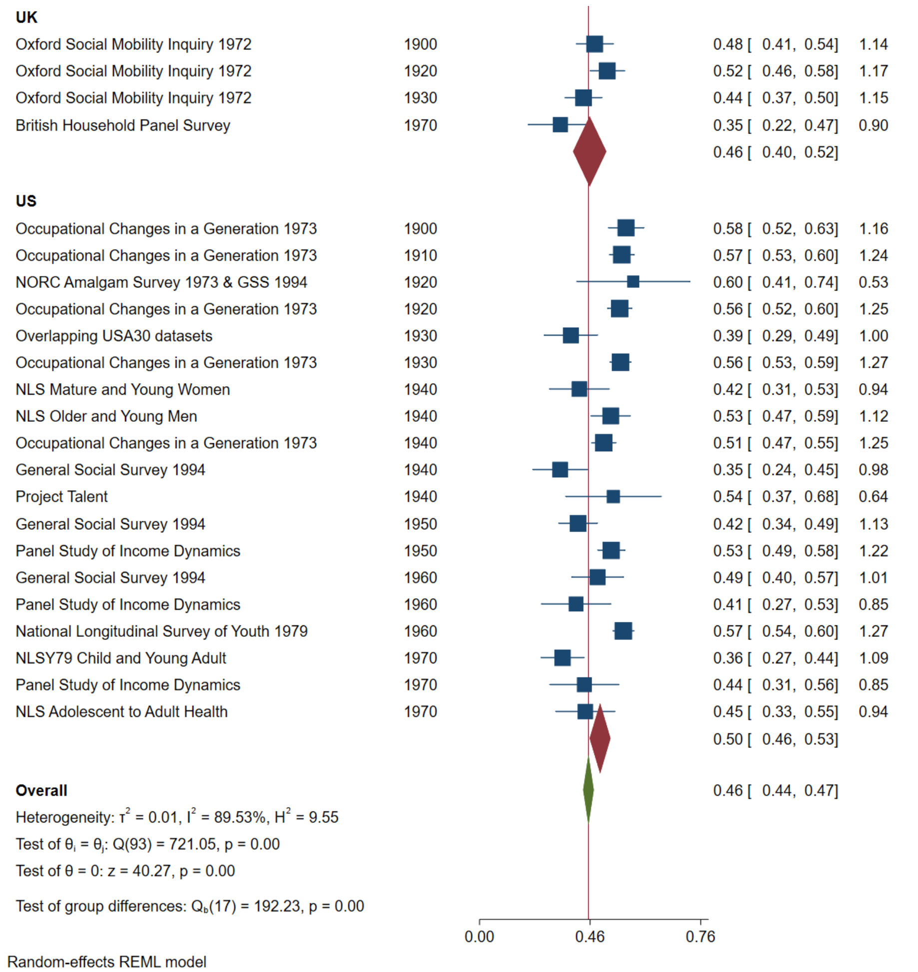

Figure S4: Forest plot of single estimated sibling correlations per sample, by country *(continued)*

H Coefficients from meta-regression of single estimated sibling correlations per sample on country and cohort

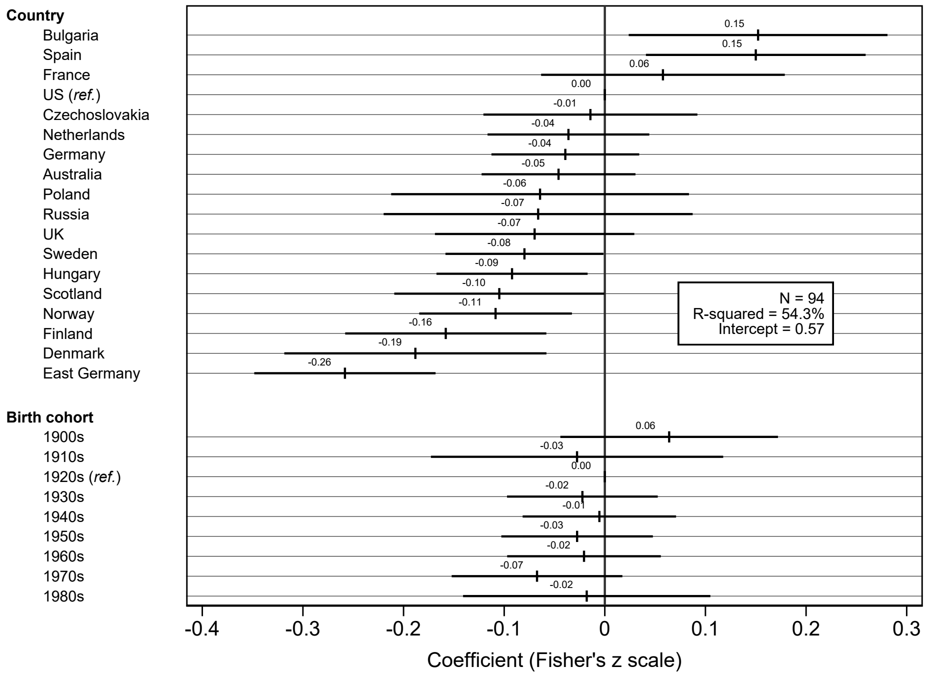

Figure S5: Coefficients from meta-regression of single estimated sibling correlations per sample on country and cohort

I Predicted sibling correlations by country, excluding cohorts born 1920 and earlier

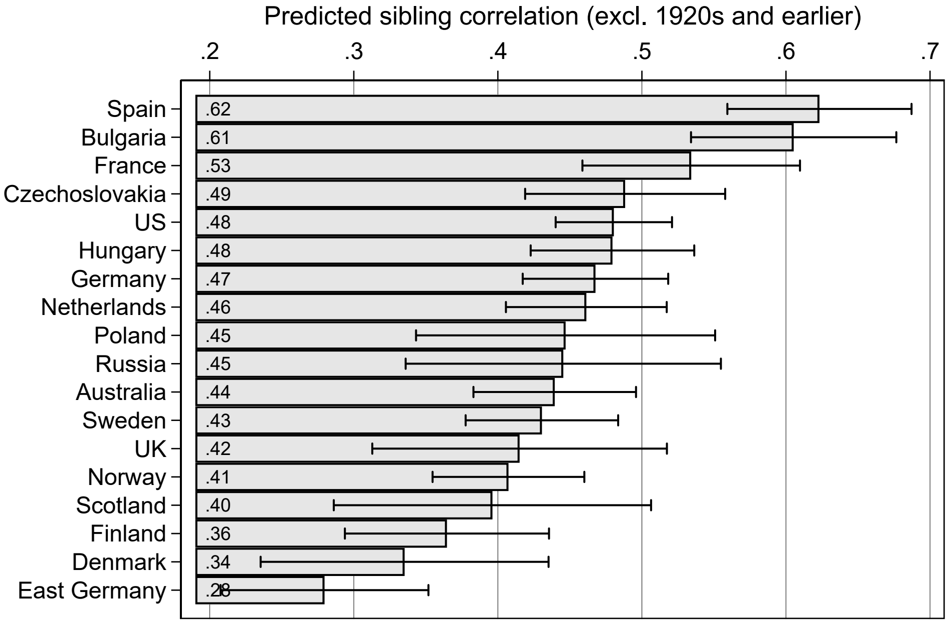

Figure S6: Predicted sibling correlations by country as in Figure 3, but excluding cohorts born 1920 and earlier

J Predicted sibling correlations by country, based only on brother and sister correlations

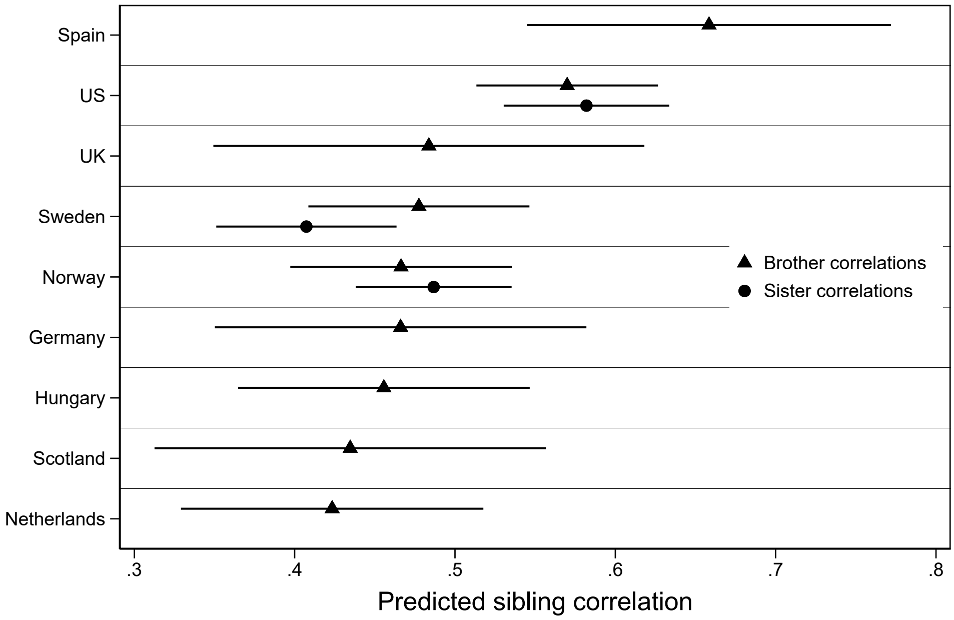

Figure S7: Predicted sibling correlations by country as in Figure 3, but based only on brother and sister correlations.

Notes: The differences from Figure 3 are 1) the estimated correlations shown are based only on observed brother correlations and sister correlations, respectively; 2) only correlations from 1930s to 1960s birth cohorts are used in the estimation, since these comprise the region of common support between the US and Nordic countries (cf. Table S1); 3) *model* is set to *variance decomposition including singletons* rather than *Pearson's r* because our data include *Pearson's r* sister correlations for Germany only, meaning that with model set to *Pearson's r*, a sister correlation could only be estimated for Germany.

K Coverage of all-siblings, brother, and sister correlations across country–cohort contexts

Table S1: Availability of all-siblings (A), brother (B), and sister (S) correlations across country–cohort contexts

Country	1900	1910	1920	1930	1940	1950	1960	1970	1980
Australia			A	A	A	A	A	A	
Bulgaria					A	A	A		
Czechoslovakia				A	A	A	A		
Denmark								B, S	A, B, S
East Germany			A	A	A	A	A		
Finland						A	A	A	
France						A		A	
Germany			A	A, B, S	A, B, S	A, B, S	A	A	A
Hungary	B	A	A, B	A, B	A	A, B	A, B		
Netherlands			A	A, B	A	A, B	A, B		
Norway				A, B, S	A, B, S	A, B, S	A, B, S	A, B, S	B, S
Poland						A	A		
Russia						A	A		
Scotland	B		B	B		B			
Spain			B	A, B	A	A, B	A		
Sweden				B	A, B	A, B, S	A, B, S	A	
United Kingdom	B		B	B				A	
United States	B	B	B	A, B	A, B, S	A, B, S	A, B, S	A	

L Country–cohort meta-analytic estimates and indicators of income inequality

Table S2: Country–cohort meta-analytic estimates and indicators of income inequality

			Income inequality	
Country	Birth cohort	Sibling corr.	Gini	Top 10 % share
Australia	1920	0.57	.	0.34
Australia	1930	0.50	.	0.34
Australia	1940	0.48	.	0.28
Australia	1950	0.46	0.26	0.28
Australia	1960	0.33	0.30	0.25
Australia	1970	0.35	0.32	0.26
Bulgaria	1940	0.44	.	.
Bulgaria	1950	0.59	0.25	.
Bulgaria	1960	0.67	0.25	.

Continued on next page

Continued from previous page

Country	Birth cohort	Sibling corr.	Gini	Top 10 % share
Czechoslovakia	1930	0.12	.	.
Czechoslovakia	1940	0.52	0.27	.
Czechoslovakia	1950	0.49	0.23	.
Czechoslovakia	1960	0.52	0.21	.
Denmark	1970	0.29	0.26	0.27
Denmark	1980	0.35	0.24	0.28
East Germany	1920	0.26	.	.
East Germany	1930	0.29	.	.
East Germany	1940	0.30	.	.
East Germany	1950	0.28	.	.
East Germany	1960	0.26	.	.
Finland	1950	0.38	0.33	.
Finland	1960	0.33	0.27	.
Finland	1970	0.36	0.23	0.26
France	1950	0.51	0.42	0.37
France	1970	0.53	0.34	0.32
Germany	1920	0.44	.	0.33
Germany	1930	0.48	.	.
Germany	1940	0.47	.	0.31
Germany	1950	0.42	.	0.28
Germany	1960	0.50	0.29	0.28
Germany	1970	0.54	0.28	0.29
Germany	1980	0.49	0.29	0.31
Hungary	1900	0.41	.	.
Hungary	1910	0.16	.	.
Hungary	1920	0.34	.	.
Hungary	1930	0.40	.	.
Hungary	1940	0.51	.	0.22
Hungary	1950	0.49	0.21	0.19
Hungary	1960	0.54	0.19	0.19
Netherlands	1920	0.51	.	0.38
Netherlands	1930	0.44	.	0.39
Netherlands	1940	0.54	.	0.32
Netherlands	1950	0.46	0.32	0.30
Netherlands	1960	0.41	0.28	0.26
Norway	1930	0.40	.	0.36
Norway	1940	0.42	.	0.35
Norway	1950	0.42	0.24	0.33

Continued on next page

Continued from previous page

Country	Birth cohort	Sibling corr.	Gini	Top 10 % share
Norway	1960	0.41	0.28	0.30
Norway	1970	0.40	0.25	0.25
Norway	1980	0.39	0.26	0.27
Poland	1950	0.41	0.26	.
Poland	1960	0.47	0.25	.
Russia	1950	0.49	.	0.24
Russia	1960	0.39	.	0.22
Scotland	1900	0.55	.	.
Scotland	1920	0.39	.	.
Scotland	1930	0.39	.	.
Scotland	1950	0.42	.	.
Spain	1920	0.38	.	.
Spain	1930	0.64	.	.
Spain	1940	0.63	.	.
Spain	1950	0.57	0.37	.
Spain	1960	0.68	0.35	.
Sweden	1930	0.45	.	0.37
Sweden	1940	0.45	.	0.32
Sweden	1950	0.41	0.31	0.33
Sweden	1960	0.40	0.24	0.29
Sweden	1970	0.44	0.23	0.27
UK	1900	0.48	.	0.35
UK	1920	0.52	.	0.35
UK	1930	0.44	.	0.30
UK	1970	0.35	0.33	0.30
US	1900	0.58	.	0.44
US	1910	0.57	.	0.46
US	1920	0.56	.	0.48
US	1930	0.49	0.37	0.40
US	1940	0.47	0.36	0.36
US	1950	0.48	0.36	0.36
US	1960	0.51	0.36	0.34
US	1970	0.40	0.36	0.36

M References Supplementary Materials

Anderson, Lewis R., Patrick Präg, Evelina T. Akimova, and Christiaan Monden, 2024. 'Replication Materials to: The Total Effect of Social Origins on Educational Attainment. Meta-Analysis of Sibling Correlations From 18 Countries.' *Open Science Framework* doi: 10.17605/OSF.IO/XHCWR.

Björklund, Anders and Kjell G. Salvanes, 2011. 'Education and Family Background. Mechanisms and Policies.' In Hanushek, Eric A., Stephen Machin, and Ludger Woessmann, eds., *Handbook of the Economics of Education*, volume 3, pp. 201–247. Amsterdam: Elsevier. doi: 10.1016/B978-0-444-53429-3.00003-X.

Black, Sandra E. and Paul J. Devereux, 2011. 'Recent Developments in Intergenerational Mobility.' In Card, David and Orley Ashenfelter, eds., *Handbook of Labor Economics*, volume 4B, pp. 1487–1541. Amsterdam: Elsevier. doi: 10.1016/S0169-7218(11)02414-2.

Branigan, Amelia R., Kenneth J. McCallum, and Jeremy Freese, 2013. 'Variation in the Heritability of Educational Attainment. An International Meta-Analysis.' *Social Forces* 92(1): 109–140. doi: 10.1093/sf/sot076.

Breen, Richard and John Ermisch, 2021. 'Sibling Models, Categorical Outcomes, and the Intra-Class Correlation.' *European Sociological Review* 37(3): 497–504. doi: 10.1093/esr/jcaa057.

Cawley, John, Euna Han, Jiyoon Kim, and Edward C. Norton, 2023. 'Genetic Nurture in Educational Attainment.' *Economics & Human Biology* 49(101239): 1–13. doi: 10.1016/j.ehb.2023.101239.

Chamberlain, Gary and Zvi Griliches, 1975. 'Unobservables with a Variance Components Structure. Ability, Schooling, and the Economic Success of Brothers.' *International Economic Review* 16(2): 422–49. doi: 10.2307/2525824.

Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan, 1972. *Socioeconomic Background and Achievement*. New York: Seminar Press.

Griliches, Zvi, 1979. 'Sibling Models and Data in Economics. Beginnings of a Survey.' *Journal of Political Economy* 87(5, Part 2): S37–S64. doi: 10.1086/260822.

Lecavelier des Etangs-Levallois, Céline and Arnaud Lefranc, 2015. *Sibling Correlations in Terms of Education, Profession, and Earnings in France*. Cergy-Pontoise: THEMA, Université de Cergy-Pontoise.

Mayer, Karl-Ulrich, 2015. 'The German Life History Study. An Introduction.' *European Sociological Review* 31(2): 137–143. doi: 10.1093/esr/jcv011.

Mazumder, Bhashkar and David I. Levine, 2004. 'The Growing Importance of Family and Community: An Analysis of Changes in the Sibling Correlation in Earnings.' *Federal Reserve Bank of Chicago Working Paper* WP 2003–24.

Müller, Walter, 1972. 'Family Background, Education, and Career Mobility.' *Social Science Information* 11(5): 223–255. doi: 10.1177/053901847201100510.

Smits, Jeroen and Christiaan Monden, 2011. 'Twinning across the Developing World.' *Plos One* 6(9): e25239. doi: 10.1371/journal.pone.0025239.

Wells, Thomas, 1995. 'Does Family Background Affect Educational Attainment Differently According to Family Structure, Birth Order, and Sex?' *NSFH Working Paper* 70.

Author affiliations

Lewis R. Anderson
Department of Social Policy and Intervention, Institute for New Economic Thinking at the Oxford Martin School, University of Oxford

Patrick Präg
Center for Research in Economics and Statistics (CREST), École nationale de la statistique et de l'administration économique (ENSAE), Institut Polytechnique de Paris

Evelina T. Akimova
Department of Sociology, Purdue University; Leverhulme Center for Demographic Science, Nuffield College, University of Oxford

Christiaan Monden
Nuffield College, Department of Sociology, University of Oxford