Heritability of labor market outcomes:

Evidence from the population of twins in the United States

David Card
UC Berkeley

Jesse Rothstein
UC Berkeley

Moises Yi
US Census Bureau

March 2026

DRAFT - PLEASE DO NOT CIRCULATE

Abstract

We use data on over 40,000 twin pairs observed as teenagers in the 2000 Census, linked to their earnings roughly 20 years later in the Longitudinal Employer Household Dynamics data file, to study the heritability of labor market outcomes. Traditional twins studies identify heritability from the contrast between identical (monozygotic) and fraternal (dizygotic) twins. We extend the classical twins design to contrasts between same-sex twins, some of whom are monozygotic, and opposite-sex twins, all of whom are dizygotic. We further extend the model to allow for unobserved family-by-gender influences on labor market outcomes. The extended model identifies heritability from the difference-in-differences of same-sex vs. opposite-sex twins vs. siblings. We find evidence of heritability of employment and earnings, and of the two main components of earnings in an AKM (Abowd et al., 1999) decomposition: person effects that are portable across jobs; and employer-specific pay premiums. However, the apparent heritability of pay premiums is driven by the high share of same-sex twins who work at the same establishment, consistent with a social interactions channel rather than a common genetic influence. This finding suggests caution in interpreting conventional estimates of heritability for socially influenced outcomes.

Acknowledgments

We are grateful to Nicole Gandre and Rina Nagashima for outstanding research assistance, and to the Russell Sage Foundation for funding. We thank Silvia Barcellos, David Cesarini, Sasha Gusev, Paige Harden, and Patrick Turley for helpful comments and discussions. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance (Project 6000266: CBDRB-FY25-CES024-003, CBDRB-FY25-CES011-002, CBDRB-FY26-CES011-001, CBDRB-FY26-CES011-005, CBDRB-FY26-CES011-007).

Significance Statement

Why do families matter so much for the labor market success of their children? One explanation is shared hereditary factors; another is family resources and family-related preferences. We use data for twins to measure the share of the variation in earnings attributable to inherited genetic factors versus other shared family characteristics and purely idiosyncratic factors. Standard methods indicate substantial genetic heritability of earnings, and of both individual and firm premium components. But when pairs who work at the exact same firm are excluded, firm effects are no longer heritable. This suggests caution in interpreting heritability estimates for socially determined outcomes, where a tendency of twins to copy each other can inflate measured heritability.

Introduction

Social scientists have long been interested in the intergenerational transmission of economic status, as measured by occupation (Blau and Duncan, 1967; Jencks, et al., 1979; Goldthorpe, et al., 1980; Hout, 1988), earnings (Behrman and Taubman, 1985; Becker and Tomes, 1986; Zimmerman, 1992), or income (Treiman and Hauser, 1977; Solon, 1992). Part of correlation in economic status between parents and children is due to parents' abilities to use their resources and social status to help their children. Part may also be due to genetic effects. Typical data sources, however, do not allow a clear distinction between these alternative channels.

Twins studies offer a solution to this problem, making it possible to measure the heritability of physical traits like height (Silventoinen et al. 2003) as well as of social outcomes like IQ (Plomin, Owen, and McGuffin, 1994), completed education (Bingley, Cappellari, and Tatsiramos, 2024), and income (Taubman, 1976; Björklund, Jäntti, and Solon, 2005; Benjamin et al., 2012; Sacerdote 2011).[1] Monozygotic (identical; MZ) twins share all of their genes, while dizygotic (fraternal; DZ) twins share some but not all genes. Both types of twin pairs share a family environment. Thus, in the standard ACE model (also known as the "classical twin design" or "CTD" – see, e.g., Visscher 2004), the heritability of a trait, h^2, defined as the share of the

[1] A separate social sciences literature uses differences between siblings or twins to identify the causal return to education, on the grounds that twins have identical ability and thus differences between them removes omitted variables bias (Griliches, 1979; Ashenfelter and Krueger, 1994; Card, 1999).

variance of the trait attributable to genes, is measured as twice the difference between the MZ and DZ correlations of that trait. Polderman et al. (2015) conduct a meta-analysis of 2,748 twins studies, classifying them by the outcome domain. Most relate to physical and mental health (e.g., cardiovascular, endocrine, reproduction, psychiatric), but Polderman et al. (2015) also identify 292 studies of cognitive traits (average h^2=0.468; SE=0.010), 56 of social interactions (average h^2=0.319; SE=0.038), and 35 of social values (average h^2=0.313; SE=0.019). Silventoinen et al. (2003) conduct twins studies of height using data from seven countries, finding values of h^2 from 0.68 to 0.85. A more recent meta-analysis of twins studies of height at age 17 (Dewau et al., 2025) finds mean h^2 of 0.77 for boys and 0.62 for girls. Twin and other genomic studies have often relied on biobank data. This limits them to unrepresentative samples and to the sets of outcomes collected by biobanks (e.g., Schoeler et al., 2023). Information on labor market outcomes is generally not available in biobanks, or is very limited. Other studies use twins who happen to appear in population surveys (e.g., the Add Health data), but these yield very small samples.

We report twins correlations for log earnings, based on the full population of twins from eight birth cohorts in the United States (N=44,500 twin pairs[2]). Our data do not distinguish MZ from DZ twins. However, we can distinguish same-gender (or same-sex, SS) from opposite-gender (OS) twins. All OS twins are DZ, while SS twins are a mix of MZ and DZ, with estimable proportions.

If boys and girls are treated differently in some families, OS twin correlations may be attenuated relative to SS DZ correlations. We extend the ACE model to allow for family-by-gender effects, which we can identify by comparing SS to OS siblings who are not twins. The SS-OS difference in correlations in earnings for twins is much larger than the SS-OS difference in correlations for siblings. In our extended ACE model, we identify h^2 from the difference-in-differences of these correlations -- SS minus OS twins vs. SS minus OS siblings. Using this approach we estimate substantial heritability of log earnings (h^2=0.36; SE=0.03).

[2] Our analyses rely on confidential data maintained by the U.S. Census Bureau. Sample sizes are rounded to the nearest 100.

There have been a few twins-based estimates of the heritability of labor market outcomes. However, our samples are much larger and more representative than others that have been used (e.g., Bingley, et al., 2024). Our specific h^2 estimates have to interpreted carefully as they rely on the assumptions of the extended ACE (E-ACE) model. Specifically, the model assumes that the family environment is equally shared for MZ as for DZ twins, and for SS vs. OS twins, net of gender differences that apply equally to SS and OS non-twin siblings (Gusev, 2024b, 2025; Bingley, et al., 2024; Burt and Simons, 2014). Moreover, the model ignores the fact that some of the shared component of earnings for identical twins may be attributable to direct interactions between the twins (Kaprio, Koskenvuo, and Rose, 1990).

We extend the analysis to examine components of individual log earnings. A burgeoning recent literature in labor economics decomposes log earnings into "person" and "firm" components, using what is known as the AKM model (Abowd, Kramarz, and Margolis, 1999; see also Card, Heining, and Kline, 2013 and Kline, 2024). The "person effect" is the permanent component of a person's earnings that is stable as the person switches across different jobs, and is often interpreted as reflecting the person's human capital (broadly interpreted to include personality traits, ambition, etc.). The "firm effect" is the common component of earnings for all workers at the same firm, net of their person effects, and is often interpreted as reflecting a firm's pay policies. In standard economic models of the labor market, all firms should pay the same wage to a given worker, but studies typically find that some firms pay higher (or lower) wages relative to others, accounting for 10-20% of the variation in wages (e.g., Card, Heining, and Kline 2013). Of great interest is understanding which workers obtain jobs at high-paying firms, and why (e.g., Card, Rothstein, and Yi 2024b, 2026; Card, Cardoso, and Kline 2016; Gerard et al., 2021). We use the AKM model to decompose earnings into individual and firm effects, and estimate heritability of each.

Our E-ACE model implies strong heritability of both person (h^2=0.40; SE=0.03) and firm (h^2=0.21; SE=0.03) effects. That the person component is heritable is unsurprising, as it presumably reflects worker abilities and skills (e.g., education) that were already known to be heritable. It is more surprising that the firm component is heritable, as the part of earning that is not due to human capital is often interpreted as "luck" (Becker and Tomes, 1979). We

document heritability of firm effects even within regions and within industries, suggesting that it is not solely due to inherited preferences over geography or type of work.

Economic outcomes like earnings are different from physical outcomes like height or IQ, in that they inherently depend on social interactions. For example, whether an individual winds up at a particular job depends on how much they value specific features of the job – such as who else works there. Such social interactions could inflate the correlation between MZ twins relative to the correlation of siblings. In the language of Manski (1993), the objects of interest in CTD studies are "correlated effects" – the determinants of the outcomes of a twin pair that reflect the influence of unobserved but common genes or family environments. But there may also be what Manski identified as "endogenous effects." Suppose, for example, that same-sex twins have a stronger preference to work together at the same job than do same-sex siblings. This will lead to an upward bias in the estimated heritability of firm-specific pay premiums from the CTD. In our data, fully 11.4% of SS twin pairs are observed working for the same firm, nearly twice the rate for non-twin SS siblings (5.7%) and triple or more the rate for OS twins (3.8%) or siblings (2.2%).

It is unlikely that the decision to work at the same firm reflects genetic effects; it is much more likely to be an endogenous social effect. Because siblings working at the same firm necessarily have the same firm effects, the high rate of overlap in firm choice among SS twins inflates the estimated heritability of these effects. To address this, we re-estimate heritability of our outcomes in a sample excluding sibling pairs who work at the same firm. The estimated heritability of firm effects falls substantially and is no longer statistically distinguishable from zero (h^2=0.07; SE=0.08). This points to the presence of endogenous interactions among twins (e.g., they choose their jobs together) that might otherwise be interpreted as genetic heritability in standard models. Such endogenous interactions could also lead to inflated estimates of the heritability of other social outcomes (for example, college attainment).

Results

The linked data, described in the supplemental materials, include 44,500 twin pairs with sufficient labor market attachment to measure their wages. Of these, 31,500 are SS and 13,000

are OS. The data also include 507,000 SS sibling pairs and 500,000 OS sibling pairs. Among children (including non-twins) born in the same cohorts and observed at the same dates 50.8% are male. We take that as an estimate of the fraction male among all births meeting the conditions for our sample inclusion (including surviving to adulthood and having observed earnings). Using the formula presented in the methods section, the implied share of SS twins who are MZ is s=58.7%.

Figure 1 presents the correlations of various outcomes between SS twins (ρ_{SS}), OS twins (ρ_{OS}), SS siblings (ρ_{SSsib}), and OS siblings (ρ_{OSsib}). The first panel shows the correlations of log average quarterly earnings. This is 0.543 (SE=0.004) for SS twins, and much lower – between 0.265 and 0.386 -- for the other comparisons. SS siblings have a stronger correlation than OS siblings, consistent with gender-specific "common" family environments, but the SS-OS difference is much larger for twins. The difference-in-differences is $\left((\rho_{SS} - \rho_{OS}) - \left(\rho_{SS}^{sib} - \rho_{OS}^{sib}\right)\right)$=0.105 (SE=0.009). By our extended ACE model, developed in the methods section, this implies h^2=0.36 (SE=0.03).

The next two panels separate log average quarterly earnings into person effect and firm effect components. In each case, ρ_{SS} is much larger than the other three correlations, and the difference-in-differences is positive – more so for the person effect (h^2=0.40, SE=0.03) than for the firm effect (h^2=0.21, SE=0.03). The remainder of the figure decomposes the firm effects into three components: A location premium, defined as the average firm effect among all firms in the commuting zone (Card, Rothstein, and Yi, 2025); an industry premium, defined as the average firm effect among all firms in the industry (Card, Rothstein, and Yi, 2024a); and a remainder. All three are correlated among twins and siblings, with the implied heritability positive for each.

These estimates all condition on labor force attachment – pairs in which either sibling does not work are excluded. The remaining panels in Figure 1 present sibling and twin correlations of three measures of employment status, estimated over the full population of sibling pairs. All show modest positive correlations among siblings, stronger for SS twins than for any of the other groups.

Table 1 presents the correlations plotted in Figure 1, along with differences between SS and OS pairs, the difference-in-differences, and the implied heritability. All outcomes show positive heritability, with larger estimates for overall log earnings and the person effect than for others. All except the location premium, which is imprecisely estimated due to the large share of children who remain in the CZ where they grew up, are significantly different from zero. Heritability estimates are plotted in **Figure 2**.

Figure 3 presents ρ_{SSsib} and ρ_{OSsib} computed over subsamples of siblings separated by different gaps in age, ranging from one year to three years. Among non-twin siblings, correlations decay as the age gap grows, but not quickly. This suggests that common family environmental influences on earnings (the C component of the ACE model) do not change very dramatically over time. The SS and OS correlations decay at similar rates, supporting our modeling assumption that gender differences are additive with common family environment differences. A test of the hypothesis that the SS-OS difference in log earnings correlations is equal across all age gaps from 4 to 12 quarters fails to reject (χ^2=4.21, d.f.=8, p=0.84). However, when we consider the twins (shown in Figure 3 with an age gap of 0), we see higher SS correlations for nearly all outcomes than for SS siblings who are even one year apart. The twin-sibling difference is smaller for OS twins and siblings. This buttresses the interpretation that our difference-in-differences estimates reflect the influence of common genes rather than drift in family environments across siblings separated in age.

Table 2 presents estimates of h^2 for two samples; these are also displayed in **Figure 4**. The first column repeats estimates for our main sample from Table 1. Column 2 excludes twin and sibling pairs who work at the same firm. SS twins have a heightened tendency to work at the same firm: 11.4% of SS twins are observed working at the same firm, as compared with 5.7% of SS siblings, 3.8% of OS twins, and 2.2% of OS siblings. The heritability estimates for log earnings and for the AKM person effect are similar when same-firm pairs are excluded, but the estimated heritability of the AKM firm effect is much reduced (h^2=0.07, SE=0.08) and no longer significant.

Discussion

Interpretation of CTD estimates is complex, especially for social outcomes. The estimated heritability of outcomes may be biased in the presence of assortative mating, genetic dominance, population stratification, or violations of the assumption that the home environment is shared to the same degree for MZ as for DZ twins. Moreover, even when estimated without bias, heritability need not indicate direct genetic effects – it also may reflect indirect or cultural effects (e.g., effects of the parents' genes on a child's outcomes, not operating through the child's own genes; see, e.g., Gusev (2024a; 2025), Howe et al. (2022)). In our setting, another channel arises that may not be relevant for outcomes like height or IQ: Twins may influence each other, inflating the correlation in their outcomes.

We find that labor market outcomes are heritable, with estimated heritability ranging from 0.18 (not significant) for the location premium component of earnings – interpreted as the extent to which people live in high-wage locations – to 0.40 for the person effect component. Heritability of employment is between these, ranging from 0.23 to 0.32 depending on the specific measure.

When we exclude sibling pairs working at the same firm, to limit the effect of social interactions between twins, the overall earnings and person effect heritability are not greatly changed. However, the heritability of firm effects and their components falls and generally ceases to be significant. (The industry premium component remains significantly heritable, though this turns negative when we exclude siblings working in the same industry – perhaps a consequence of the limited number of industries in our sample.)

There are obvious non-genetic reasons for twins to tend to work at the same firm. Thus, we interpret the estimates excluding same-firm pairs as the best estimates of heritability. These indicate that log wages and the person effect component are strongly heritable (h^2=0.31 and 0.35 respectively) but that the firm effect (h^2=0.07) is not. The latter suggests that firm effects arise from factors that are equally shared by MZ twins and same sex siblings, like the proximity of higher-paying employers to the neighborhood where they were raised. The fall in the heritability of firm effects when twins who work for the same firm are excluded also raises a caution about potential confounding of the CTD when applied to other outcomes where individual choice plays a role. One can easily imagine similar mechanisms applying to other

outcomes that have been studied using the CTD – for example, SS twins might make similar educational choices out of a desire to stay together rather than due to shared genetic tendencies. This suggests that even our estimates of the heritability of employment, earnings, and the person effect may be inflated by unobserved social effects.

MATERIALS AND METHODS

Data

Our analysis is based on children identified in the U.S. 2000 census of the population who were the biological child of the household head. We identify twins as children in the same household with the same birthdate, to within 2 days, and siblings as children with different birthdates (nine months or more apart). The U.S. Census Bureau assigns Personal Identification Keys (PIKs) to all respondents using name and birthdate information. We limit attention twin or sibling pairs in which both children were assigned PIKs and both were between 9 and 17 years old at the time of the census. Families with more than two children can contribute multiple pairs to the sample.

These individuals are linked to data from the Longitudinal Employer Household Dynamics (LEHD) data, with information on employment, employer, and quarterly earnings. LEHD data are derived from employers' unemployment insurance tax payments, and cover all formal sector employment (but not self-employment). We use LEHD data from 2018-2022, excluding the second and third quarters of 2020 when the labor market was substantially disrupted by the COVID pandemic. We focus on individuals who are observed in the LEHD in this period with at least 8 quarters of observed earnings above $3,300 (approximately the minimum wage for full-time work), excluding the first and last quarter of each employment spell and any quarter in which the worker has multiple employers. With these restrictions, we interpret observed quarterly earnings as indicating full-quarter, full-time earnings. Our sample for analysis of earnings outcomes retains only sibling pairs in which both siblings meet the restrictions. Our primary outcome is the average log earnings for the individual across all quarters meeting the above restrictions.

We also measure employment – an indicator for having any employment during our period, the number of quarters with positive employment, and an indicator for having sufficient employment to enter into our AKM sample. These are measured over the full population of sibling pairs, not conditioning on sufficient observed earnings.

We consider two siblings to work at the same firm if one sibling is ever in our 2018-2022 data window observed at a firm where the other sibling is also observed in that window.[3] We use a similar definition for working in the same industry. When we examine correlations for siblings observed with different age gaps, we exclude non-twin sibling pairs with gaps smaller than one year, as these likely reflect adoptions or measurement error rather than actual sibling pairs.

Extended ACE Model

Let cov_g represent the covariance of long-term outcome Y_i between siblings of type g, where the types to be considered are MZ twins ($g = MZ$); same-sex DZ twins ($g = DZSS$); all same-sex twins, combining MZ and DZSS ($g = SS$); opposite-sex twins, who are all DZ ($g = OS$); same-sex, non-twin siblings ($g = SSsib$); and opposite-sex, non-twin siblings ($g = OSsib$). We can measure cov_{SS}, cov_{OS}, cov_{SSsib}, and cov_{OSsib} directly.

In the ACE model (e.g., Visscher 2004),long-term outcome Y_i is the sum of 3 factors: genetic endowment A_i, shared family environment $C_{f(i)}$, where $f(i)$ is an index function pointing to the family of person i, and E_i, representing non-shared environment: $Y_i = A_i + C_{f(i)} + E_i$. The heritability of the outcome is defined as $h^2 = V(A)/V(Y)$, the share of the variation that is explained by the genetic endowment.

MZ twins *i* and *j* are assumed to have identical genetic endowment $A_i = A_j$. DZ twins are assumed to have correlated genetic endowment $R_A = corr(A_i, A_j)$, with $R_A < 1$. As in other CTD studies, we assume $R_A = 0.5$, as would arise under random mating, while acknowledging that with assortative mating R_A might be larger.

[3] The LEHD assigns workers to establishments, the specific location where a worker works (e.g., a particular restaurant location), rather than to firms (e.g., McDonalds). All references to "firms" in fact refer to specific establishments.

Both MZ and DZ twins have identical shared environment, $C_{f(i)} = C_{f(j)}$, but independent non-shared environment. Assuming variances and the A-C covariance are the same for MZ and DZ twins, $cov_{MZ} = V(A) + V(C) + V(E) + 2cov(A, C)$ while $cov_{DZ} = R_A V(A) + V(C) + V(E) + 2cov(A, C)$. Thus, heritability can be estimated as $h^2 = (cov_{MZ} - cov_{DZ})/(1 - R_A)V(Y)$.

We cannot distinguish MZ from DZSS twins, only SS from OS. The SS twin covariance is

$$cov_{SS} = (s + (1 - s)R_A)V(A) + V(C) + V(E) + 2cov(A, C)$$

where $s \equiv p(MZ \mid SS)$ is the share of SS twins who are MZ. The OS twin covariance is the same as the DZ covariance,

$$cov_{OS} = R_A V(A) + V(C) + V(E) + 2cov(A, C).$$

This implies $h^2 = (cov_{SS} - cov_{OS})/s(1 - R_A)V(Y)$. We discuss estimation of s below; for now, we treat it as known.

We do not believe the SS-OS difference can be used to identify h^2 for labor market outcomes, however, as we expect the family environment relevant for labor market outcomes may differ for boys and girls in the same family. Some families are gender egalitarian, promoting sons' and daughters' labor market success to a similar degree, while others take more traditional views of gender roles, and may promote sons' success at the expense of daughters'. This would violate the equal environments assumption and reduce the OS correlation relative to the SS correlation.

We extend the ACE model by adding a second family environment term, G, that varies by child's gender $g(i)$: $Y_i = A_i + C_{f(i)} + G_{f(i)g(i)} + E_i$. We assume that for two siblings *i* and *j* in the same family f, $corr\left(G_{fg(i)}, G_{fg(j)}\right) = 1$ if $g(i) = g(j)$ but $corr\left(G_{fg(i)}, G_{fg(j)}\right) = \tau < 1$ if $g(i) \neq g(j)$. For simplicity, we assume $cov(A, G) = cov(C, G) = 0$, though this is not essential. In this extended model,

$$cov_{MZ} = V(A) + V(C) + V(G) + V(E) + 2cov(A, C)$$
$$cov_{DZSS} = R_A V(A) + V(C) + V(G) + V(E) + 2cov(A, C)$$
$$cov_{SS} = (s + (1 - s)R_A)V(A) + V(C) + V(G) + V(E) + 2cov(A, C)$$
$$cov_{OS} = R_A V(A) + V(C) + \tau V(G) + V(E) + 2cov(A, C)$$

In the presence of a family component that varies by gender, we cannot isolate $V(A)$ based only on the difference in covariances between SS and OS twins:

$$cov_{SS} - cov_{OS} = s(1 - R_A)V(A) + (1 - \tau)V(G).$$

To proceed, we augment this system with covariances for non-twin siblings. We assume the G correlations are the same as for twins – 1 if same-sex and τ if opposite sex. But the common family environment effect C may differ for siblings, whose time in the home only partly overlaps. We assume that the correlation of this component is $\theta < 1$ for non-twin siblings. We further assume that A is correlated the same for siblings as for DZ twins, R_A. This implies

$$cov_{SSsib} = R_A V(A) + \theta V(C) + V(G) + V(E) + 2cov(A, C)$$
$$cov_{OSsib} = R_A V(A) + \theta V(C) + \tau V(G) + V(E) + 2cov(A, C)$$
$$cov_{SSsib} - cov_{OSsib} = (1 - \tau)V(G).$$

Combining the expressions for the differences in covariances between same-sex and opposite sex twins versus siblings, we obtain:

$$V(A) = \frac{(cov_{SS} - cov_{OS}) - (cov_{SSsib} - cov_{OSsib})}{s(1 - R_A)}.$$

With this in hand, we can further compute $h^2 = V(A)/V(Y)$, $cov_{MZ} = cov_{SS} + (1 - s)(1 - R_A)V(A)$, and $cov_{DZSS} = cov_{SS} - s(1 - R_A)V(A)$.

We assume $R_A = 0.5$, though explore sensitivity to higher values. We assume the variances of all elements of our extended ACE model are constant across types of siblings (twins vs. not, SS vs. OS), and report estimated correlations rather than covariances, and h^2 rather than $V(A)$.

To estimate $s \equiv p(MZ|SS)$, we first use the "excess fraction" of SS twins above the rate for DZ twins to estimate the share of MZ twins:

$$p(MZ) = \frac{p(SS) - p(SS \mid DZ)}{1 - p(SS \mid DZ)}$$

Assuming that the sexes of DZ twins are determined independently with probability π of being male (which we estimate at 0.508), we get $p(SS|DZ) = \pi^2 + (1 - \pi)^2$. Next we use Bayes law and the fact that all MZ twins are SS to obtain:

$$p(MZ|SS) = \frac{p(MZ)}{p(SS)} = \frac{p(SS) - p(SS \mid DZ)}{p(SS) \times (1 - p(SS \mid DZ))}$$

With $p(SS) = 0.708$ (in our sample), and $p(SS \mid DZ) = 0.500$, we estimate $p(MZ \mid SS) = 0.588$.

A key assumption of our E-ACE model is that the gender-specific component of family environment (G) is perfectly correlated across siblings and additively separable from the non-gendered component. This allows us to remove the bias in $cov_{SS} - cov_{OS}$ by subtracting off the analogous difference for siblings. A second assumption is that the correlation between the non-gendered family component in sibling outcomes (the "shared family environment") is invariant to the age gap between siblings. If instead the correlation decays with the age gap, but with the same rate of decay for SS and OS siblings, then our difference in differences approach remains valid providing the distributions of age gaps are similar for SS and OS siblings. The roughly parallel and quite flat profiles of the correlations in log earnings for SS and OS siblings at different age gaps shown in Figure 3 suggest that the shared decay rate assumption is approximately true – and that the rate of decay is very small.

Standard errors

The traditional formula for the standard error of a correlation is $\frac{1-\rho^2}{\sqrt{N}}$, where *N* is the number of pairs used to compute the correlation. This does not account for potential heteroskedasticity in the underlying variables, however. We instead use a regression-based approach. If β_{XY} is the coefficient of a regression of *Y* on *X*, it is related to the correlation coefficient by $\rho_{XY} = \frac{\sigma_Y}{\sigma_X}\beta_{XY}$, where σ_Y and σ_X are the standard deviations of *Y* and *X*, respectively. We thus compute the sampling variance of the correlation coefficient as $V(\rho_{XY}) = \frac{\sigma_Y^2}{\sigma_X^2}V(\beta_{XY})$, where $V(\beta_{XY})$ is the heteroskedasticity-robust variance of β_{XY} and we treat the variances of *X* and Y as known. This yields slightly larger standard errors than the traditional formula.

Our estimate of h^2 is computed from the difference-in-differences of correlation coefficients from independent samples, multiplied by a constant $\frac{1}{s(1-R_A)}$. We treat $s(1-R_A)$ as known. Thus, the variance of h^2 is the sum of the variances of the underlying correlations, multiplied by the square of the constant.

Decomposing log earnings using AKM model

We use a decomposition based on the method described by Abowd et al. (1999) to decompose quarterly earnings in the LEDH file. In brief the model can be written as:

$$y_{it} = \alpha_i + \psi_{j(i,t)} + X_{it}\beta + \xi_{it},$$

where y_{it} represents the log earnings of individual i in quarter t, α_i is a person fixed effect, $\psi_{j(i,t)}$ is a fixed effect for the establishment at which the person works in quarter t, $j(i,t)$ is an index function giving the workplace of individual i in quarter t, $X_{it}\beta$ captures effects of time-varying observable covariates, such as experience and calendar time, and ξ_{it} is assumed to be orthogonal to the sequence of establishment effects, conditional on α_i (see Card et al, 2013). In the AKM literature α_i is interpreted as an individual's earnings capacity at any job, while ψ_j is interpreted as a wage premium or discount that all workers get at establishment j relative to pay in a reference workplace (which we take to be average employer in the restaurant sector).

We estimate this model using data on all individuals in the LEHD from 2018-I to 2023-I, omitting data from 2020-II to 2020-IV during the COVID epidemic. The sample is limited to people who have at least 8 quarters with earnings over $3,300. We use only quarters when an individual is between 20 and 60 years of age, has a single employer, and is earning at least $3,800 in the quarter. The first and last quarters on a given job are also excluded. After model estimation we assign the estimated person effects to the individuals who are either twins or siblings in our analysis sample. Similarly, we assign the mean of the estimated pay premiums they earned in all the quarters they are observed satisfying the restrictions for our estimation sample. The sum of these two components accounts for 82% of the variance in person-quarter log earnings in the cohorts from which our twin sample is drawn (Card, Rothstein, and Yi, 2026). We further decompose the firm pay premiums into location, industry, and residual components, following methods developed in Card, Rothstein, and Yi (2026; see also Card, Rothstein, and Yi, 2024a, 2024b, 2025).

References

O. Ashenfelter, A. Krueger, Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review* **84**, 1157 (1994).

G. S. Becker, N. Tomes, Human capital and the rise and fall of families. *Journal of Labor Economics* **4**, S1-S39 (1986).

J. Behrman, P. Taubman, Intergenerational earnings mobility in the United States: Some estimates and a test of Becker's intergenerational endowments model. *The Review of Economics and Statistics* **67**, 144-151 (1985).

D. J. Benjamin *et al.*, The promises and pitfalls of genoeconomics. *Annu. Rev. Econ.* **4**, 627-662 (2012).

P. Bingley, L. Cappellari, K. Tatsiramos, "On the origins of socioeconomic inequalities: Evidence from twin families" (Working paper 2024-03, Luxembourg Institute of Socio-Economic Research, 2024).

A. Björklund, M. Jäntti, G. Solon, "Influences of nature and nurture on earnings variation: A report on a study of various sibling types in Sweden" in *Unequal Chances: Family Background and Economic Success*, S. Bowles, H. Gintis, M. Osborne Groves, Eds. (Princeton Univ. Press, Princeton, NJ, 2005), pp. 145-164.

P. M. Blau, O. D. Duncan, *The American Occupational Structure* (Wiley, New York, 1967).

D. Card, "The causal effect of education on earnings" in *Handbook of Labor Economics*, O. Ashenfelter, D. Card, Eds. (Elsevier, Amsterdam, 1999), vol. 3, pp. 1801-1863.

D. Card, A. R. Cardoso, P. Kline, Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics* **131**, 633-686 (2016).

D. Card, J. Heining, P. Kline, Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics* **128**, 967-1015 (2013).

D. Card, J. Rothstein, M. Yi, Industry wage differentials: A firm-based approach. *Journal of Labor Economics* **42**, S11–S59 (2024).

D. Card, J. Rothstein, M. Yi, Reassessing the spatial mismatch hypothesis. *AEA Papers and Proceedings* **114**, 221–225 (2024).

D. Card, J. Rothstein, M. Yi, Location, location, location. *American Economic Journal: Applied Economics* **17**, 297–336 (2025).

D. Card, J. Rothstein, M. Yi, "How do neighborhoods and firms affect intergenerational mobility?" (Working paper, University of California, Berkeley, 2026).

R. Dewau *et al.*, Meta-analysis of the heritability of childhood height from 560,000 pairs of relatives born between 1929 and 2004. *American Journal of Human Biology* **37**, e24188 (2025).

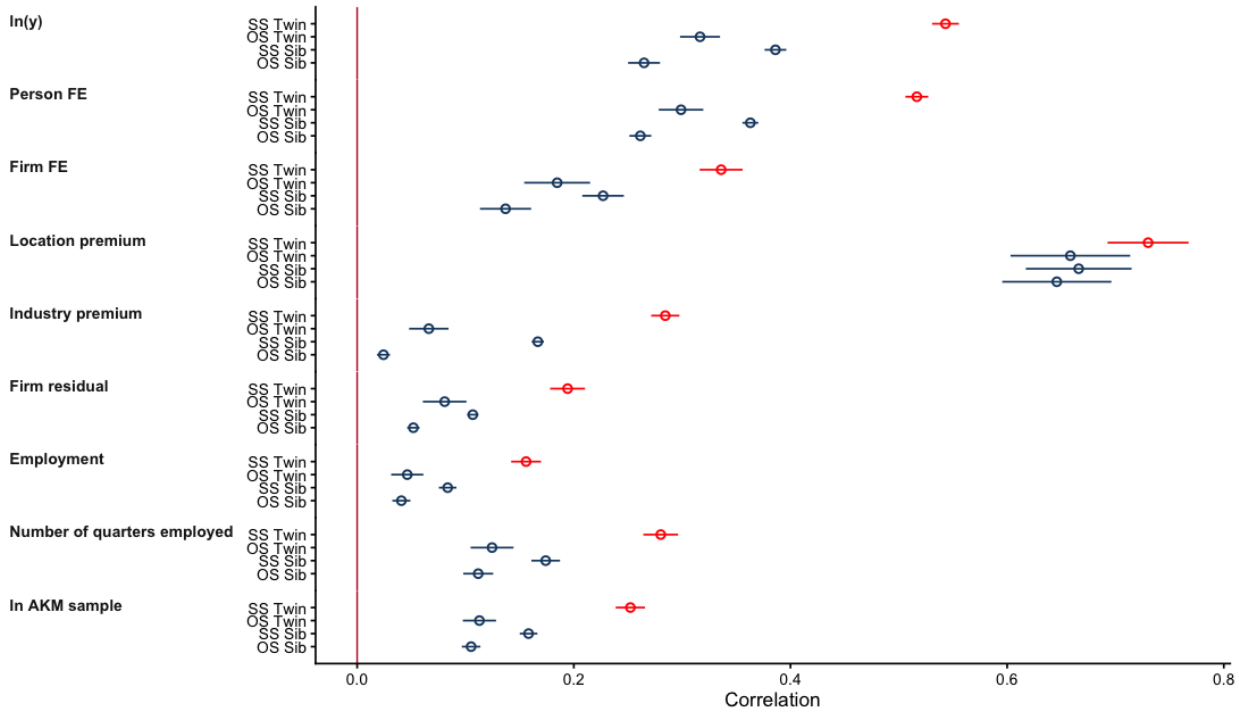
F. Gerard, L. Lagos, E. Severnini, D. Card, Assortative matching or exclusionary hiring? The impact of employment and pay policies on racial wage differences in Brazil. *American Economic Review* **111**, 3418-3457 (2021).

J. H. Goldthorpe, C. Llewellyn, C. Payne, *Social Mobility and Class Structure in Modern Britain* (Clarendon Press, Oxford, 1980).

Z. Griliches, Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy* **87**, S37-S64 (1979).

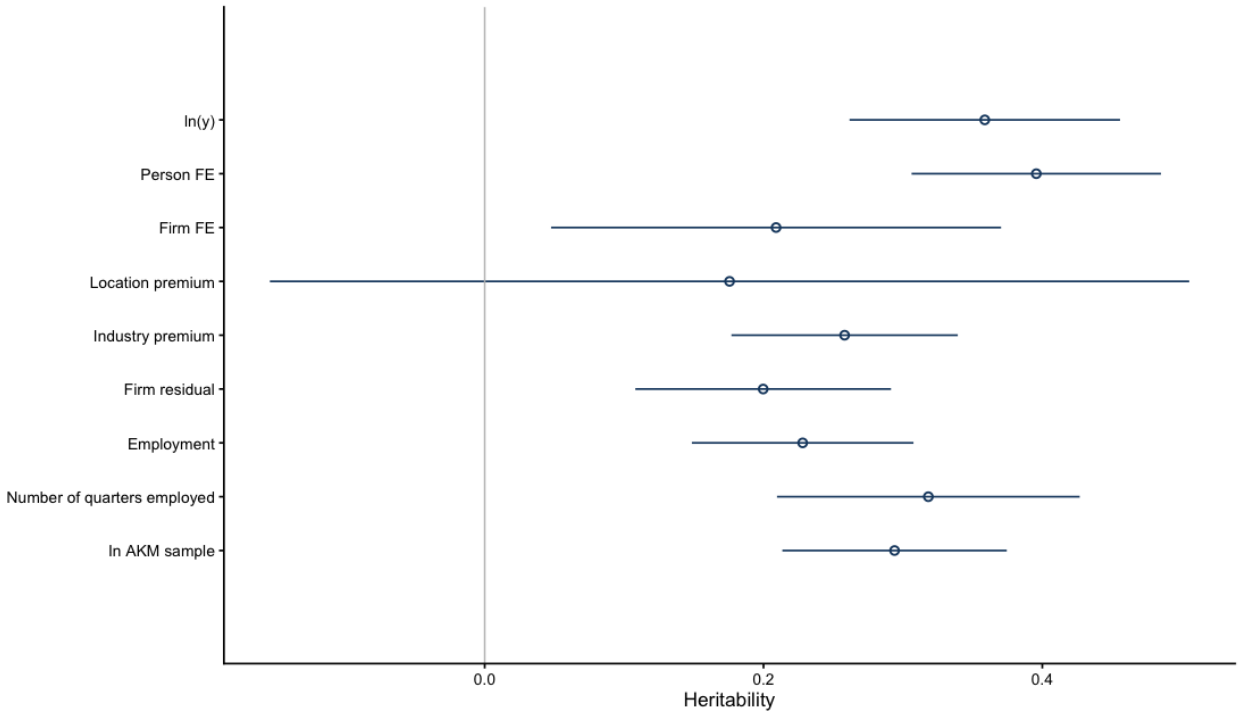
S. Gusev, No, intelligence is not like height. *The Infinitesimal* (2024). https://theinfinitesimal.substack.com/p/no-intelligence-is-not-like-height.
S. Gusev, Twin heritability models can tell you whatever you want to hear. *The Infinitesimal* (2024). https://theinfinitesimal.substack.com/p/twin-heritability-models-can-tell.
S. Gusev, The missing heritability question is now (mostly) answered. *The Infinitesimal* (2025). https://theinfinitesimal.substack.com/p/the-missing-heritability-question.
M. Hout, More universalism, less structural mobility: The American occupational structure in the 1980s. *American Journal of Sociology* **93**, 1358-1400 (1988).
L. J. Howe *et al.*, Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature Genetics* **54**, 581-592 (2022).
C. Jencks *et al.*, *Who Gets Ahead? The Determinants of Economic Success in America* (Basic Books, New York, 1979).
P. Kline, "Firm wage effects" in *Handbook of Labor Economics*, C. Dustman, T. Lemieux, Eds. (North-Holland, Amsterdam, 2024), vol. 5, pp. 115-181.
C. F. Manski, Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* **60**, 531-542 (1993).
R. Plomin, M. J. Owen, P. McGuffin, The genetic basis of complex human behaviors. *Science* **264**, 1733-1739 (1994).
T. J. C. Polderman *et al.*, Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics* **47**, 702-709 (2015).
B. Sacerdote, "Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees?" in *Handbook of Social Economics*, J. Benhabib, A. Bisin, M.O. Jackson, Eds. (North-Holland, Amsterdam, 2011), vol. 1, pp. 1-30.
T. Schoeler *et al.*, Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nature Human Behaviour* **7**, 1216-1227 (2023).
K. Silventoinen *et al.*, Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Research and Human Genetics* **6**, 399-408 (2003).
G. Solon, Intergenerational income mobility in the United States. *The American Economic Review* **82**, 393-408 (1992).
P. Taubman, The determinants of earnings: Genetics, family, and other environments; a study of white male twins. *American Economic Review* **66**, 858-870 (1976).
D. J. Treiman, R. M. Hauser, "Intergenerational transmission of income: An exercise in theory construction" in *The Process of Stratification: Trends and Analyses*, R. M. Hauser, D. L. Featherman, Eds. (Academic Press, New York, 1977), pp. 271-302.
P. M. Visscher, Power of the classical twin design revisited. *Twin Research and Human Genetics* **7**, 505-512 (2004).
D. J. Zimmerman, Regression toward mediocrity in economic stature. *The American Economic Review* **82**, 409-429 (1992).

Figure 1. Between-sibling and between-twin correlations, same-sex (SS) and opposite-sex (OS)



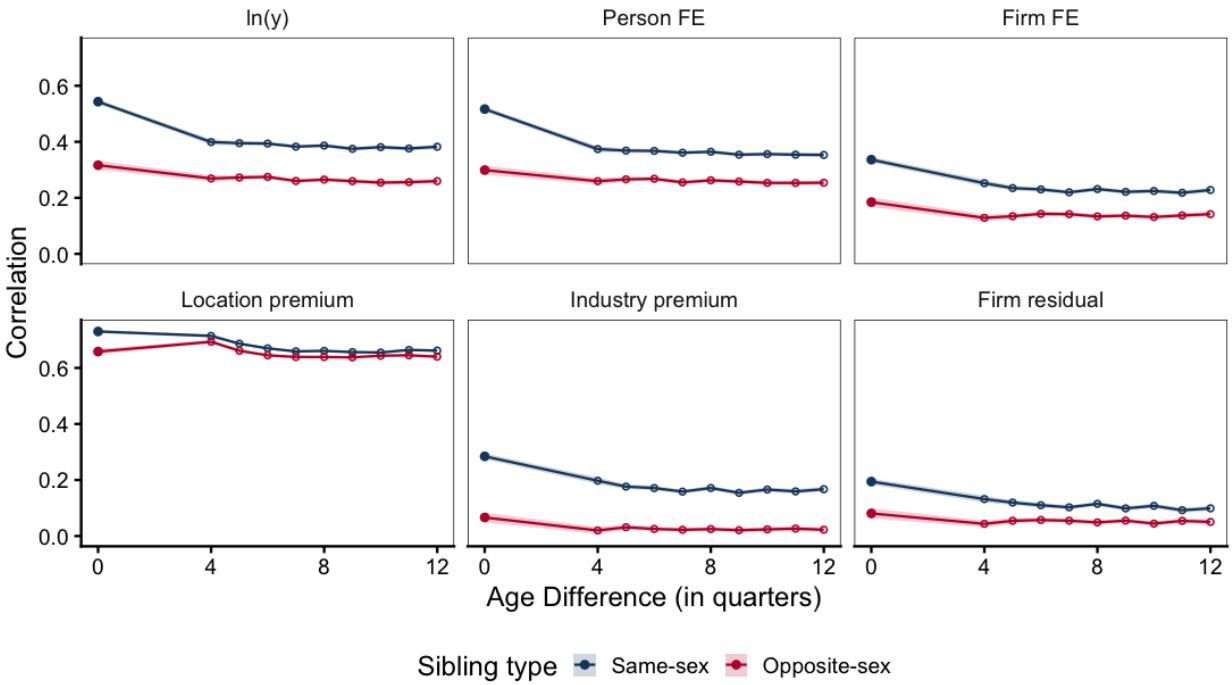
Notes: 95% confidence intervals in parentheses. Log earnings and its components, presented in the first six panels, are computed over the sample of twin and sibling pairs who are both stably employed, as described in the data section. Employment outcomes, in the final three panels, are computed over all twin pairs, including those who are not both stably employed.

Figure 2. Heritability estimates from E-ACE model



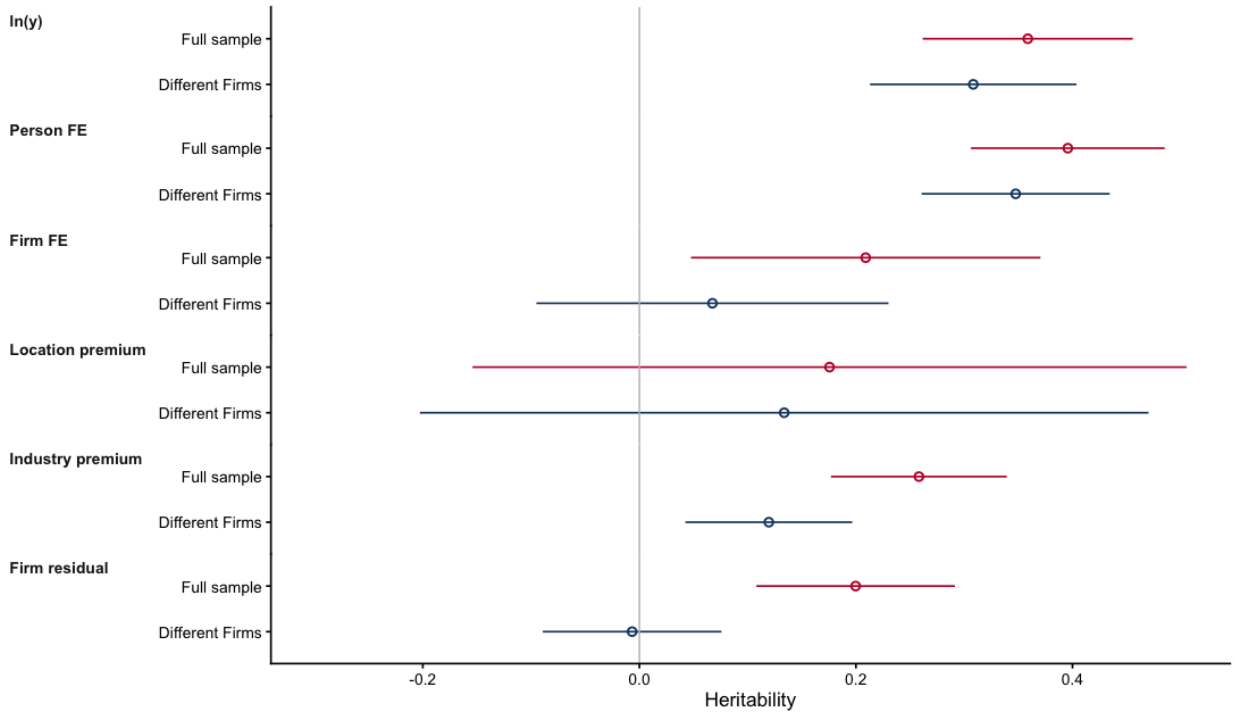
Notes: 95% confidence intervals in parentheses.

Figure 3. Between-sibling correlations, by age difference between siblings



Notes: Shaded region indicates 95% pointwise confidence interval. This is computed using traditional formula for the standard error of a correlation coefficient, pending disclosure of regression-based estimates.

Figure 4. Heritability estimates from E-ACE model, with and without sibling pairs working at the same firm



Notes: 95% confidence intervals in parentheses. “Different firms” sample is computed from sibling and twin pairs for whom the lists of firms at which the two siblings work in 2018-2023Q1 are non-overlapping.

Table 1. Correlations of outcomes among twins and siblings

	Correlations				Differences			Difference-in-differences	Implied heritability
	Twins		Siblings		SS-OS twins	SS twins-SS siblings	SS-OS siblings		
	SS	OS	SS	OS					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employment outcomes									
N (1,000s)	60	27	1,072	1,096					
Any employment	0.156	0.046	0.084	0.041	0.110	0.072	0.043	0.067	0.23
	(0.007)	(0.007)	(0.004)	(0.004)	(0.010)	(0.008)	(0.006)	(0.012)	(0.04)
No. of quarters employed	0.280	0.125	0.174	0.112	0.156	0.106	0.062	0.093	0.32
	(0.008)	(0.010)	(0.007)	(0.007)	(0.013)	(0.010)	(0.010)	(0.016)	(0.05)
In high-employment sampl	0.252	0.113	0.158	0.105	0.139	0.094	0.053	0.086	0.29
	(0.007)	(0.008)	(0.004)	(0.004)	(0.010)	(0.008)	(0.006)	(0.012)	(0.04)
Earnings outcomes									
N (1,000s)	32	13	507	500					
Log earnings	0.543	0.317	0.386	0.265	0.227	0.157	0.121	0.105	0.36
	(0.006)	(0.009)	(0.005)	(0.007)	(0.011)	(0.008)	(0.009)	(0.014)	(0.05)
AKM person effect	0.517	0.299	0.363	0.261	0.218	0.154	0.102	0.116	0.40
	(0.005)	(0.010)	(0.004)	(0.005)	(0.012)	(0.006)	(0.006)	(0.013)	(0.04)
AKM firm effect	0.336	0.185	0.227	0.137	0.151	0.109	0.090	0.061	0.21
	(0.010)	(0.015)	(0.010)	(0.012)	(0.018)	(0.014)	(0.015)	(0.024)	(0.08)
Location premium	0.730	0.658	0.666	0.646	0.072	0.064	0.020	0.052	0.18
	(0.019)	(0.028)	(0.024)	(0.025)	(0.033)	(0.031)	(0.035)	(0.048)	(0.16)
Industry premium	0.284	0.066	0.167	0.024	0.218	0.118	0.142	0.076	0.26
	(0.006)	(0.009)	(0.003)	(0.003)	(0.011)	(0.007)	(0.004)	(0.012)	(0.04)
Firm residual	0.194	0.081	0.107	0.052	0.113	0.088	0.055	0.059	0.20
	(0.008)	(0.010)	(0.003)	(0.003)	(0.013)	(0.008)	(0.004)	(0.013)	(0.05)

Notes: SS = same sex; OS = opposite sex. Standard errors in parenthesis. Correlation standard errors are computed from the regression of one sibling's outcome on the other's, with robust standard errors. Standard errors for differences, difference-in-differences, and heritability statistics are computed from the correlation standard errors using the delta method.

Table 2. Sensitivity to firm and industry overlap

	Estimated heritability	
	Full sample	No firm overlap
	(1)	(2)
Log earnings	0.36	0.31
	(0.05)	(0.05)
AKM person effect	0.40	0.35
	(0.04)	(0.05)
AKM firm effect	0.21	0.07
	(0.08)	(0.08)
Location premium	0.18	0.13
	(0.16)	(0.17)
Industry premium	0.26	0.12
	(0.04)	(0.04)
Firm residual	0.20	-0.01
	(0.05)	(0.04)
Number of twin pairs (1,000s)	45	41
Fraction in same firm	9%	0%
Comparison: Siblings		
Number of sibling pairs (1,000s)	1,007	967
Fraction in same firm	4%	0%

Notes: "No firm overlap" sample excludes twin or sibling pairs for which there is any overlap in the list of firms at which the two siblings worked in 2018-2023Q1. Standard errors in parenthesis. Standard errors for correlations between SS and OS twins and siblings are computed from the regression of one twin or sibling's outcome on the other's, with robust standard errors. Standard errors for heritability statistics are computed from the correlation standard errors using the delta method.