

Socioeconomic status and genetic influences on cognitive development

David N. Figlio^a, Jeremy Freese^{b,1}, Krzysztof Karbownik^c, and Jeffrey Roth^d

^aSchool of Education and Social Policy, Northwestern University, Evanston, IL 60208; ^bDepartment of Sociology, Stanford University, Stanford, CA 94305; ^cInstitute for Policy Research, Northwestern University, Evanston, IL 60208; and ^dDepartment of Pediatrics, University of Florida, Gainesville, FL 32608

Edited by James J. Heckman, The University of Chicago, Chicago, IL, and approved October 18, 2017 (received for review May 26, 2017)

Accurate understanding of environmental moderation of genetic influences is vital to advancing the science of cognitive development as well as for designing interventions. One widely reported idea is increasing genetic influence on cognition for children raised in higher socioeconomic status (SES) families, including recent proposals that the pattern is a particularly US phenomenon. We used matched birth and school records from Florida siblings and twins born in 1994–2002 to provide the largest, most population-diverse consideration of this hypothesis to date. We found no evidence of SES moderation of genetic influence on test scores, suggesting that articulating gene-environment interactions for cognition is more complex and elusive than previously supposed.

socioeconomic status | cognition | behavior genetics | environmental moderation | twin studies

That genes and environments combine to influence cognitive development is broadly recognized, yet clear specifications of how they combine remain elusive. Behavioral geneticists have found that genetic differences are more influential on cognition for persons raised in more advantaged environments, a result sometimes called the Scarr-Rowe interaction (1). The animating idea is that social disadvantage compromises the extent to which a child's genetic potential is realized. As a result, the ultimate influence of genetic endowment is lower in these environments, which also implies that higher heritability estimates reflect improved social conditions (2–4).

Although there have been striking findings supporting the hypothesis (5–10), results are inconsistent (11–15), and a recent meta-analysis indicated only modest support (16). Notably also, the meta-analysis found that the hypothesis fared much better in studies of US samples than samples elsewhere, mainly Northern/Western Europe and Australia.

Potential explanations of this divergence include the possibilities that socioeconomic variation in the United States is simply larger in magnitude or that the US educational system is less effective at helping disadvantaged students reach their potential (16). Child poverty rates, rates of children living in homes with deprived educational resources, and inequality in educational achievement are all higher in the United States than in countries in which null results have been reported (17, 18). If heritability of cognition reflects opportunity, then differential heritability by socioeconomic status (SES) in the United States could be interpreted as a consequence of some combination of social disparities in the United States. Of course, any such conclusion is predicated on establishing that results really are different for the United States.

Current Study

This study considered the hypothesis using unique confidential population-level administrative data that match birth and public school records for all Florida children born between 1994 and 2002. Birth records were matched to school records on the basis of names, birth dates, and social security numbers. The rate of these matches is consistent with expectations from US Census data about the percentage of children born in Florida who

subsequently attend Florida public schools. Our investigations indicate that rates of being able to match records for one twin but not the other are extremely low and unlikely to bias findings (19). By using state records, we were able to assemble a sample that has longitudinal test score information and is an order of magnitude larger than prior studies (24,640 twins with matched birth-school records and test score information).

Our method does not depend on locating, recruiting, and retaining twins to participate in data collection efforts, which is important because this often considerably reduces the representation of twins from disadvantaged backgrounds in study samples. Having data from Florida allowed us to represent a broader range of socioeconomic backgrounds better than past work, because the state has comparable or greater socioeconomic inequality than any other population for which the hypothesis has been considered so far. In our twin sample, 25.6% of mothers were African American, 18.0% were Hispanic, 14.5% were aged ≤ 21 y, and 32.0% were unmarried at time of birth. Although an earlier survey-based study in Florida children found evidence of SES moderation of the heritability of achievement test scores, that sample had only 577 twin pairs and measured the SES of schools rather than families themselves (20).

Twin studies in behavioral genetics typically identify the genetic contribution to an outcome's variance using the difference in genetic relatedness between monozygotic (MZ; "identical") and dizygotic (DZ; "fraternal") twins. Trait heritability is then a function of the intraclass correlations (ICCs; the ratio of between-pair variance to total variance) between samples of MZ and DZ twins drawn from the same population (21). A hypothesis that SES moderates heritability implies that the relationship between ICCs differs between SES groups.

Administrative data generally do not contain information on zygosity. However, because all opposite-sex (OS) twin pairs are

Significance

A prominent hypothesis in the study of intelligence is that genetic influences on cognitive abilities are larger for children raised in more advantaged environments. Evidence to date has been mixed, with some indication that the hypothesized pattern may hold in the United States but not elsewhere. We conducted the largest study to date using matched birth and school administrative records from the socioeconomically diverse state of Florida, and we did not find evidence for the hypothesis.

Author contributions: D.N.F., J.F., K.K., and J.R. designed research; D.N.F., J.F., and K.K. performed research; J.F. and K.K. analyzed data; and D.N.F., J.F., K.K., and J.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence should be addressed. Email: jfreese@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1708491114/-DCSupplemental.

DZ whereas same-sex (SS) twin pairs contain approximately a 50–50 mix of MZ and DZ twin pairs, SS twin pairs will, on average, be genetically more similar than OS twin pairs. One remaining complication with using administrative data is that among DZ twins, SS twins could also be more similar to one another on an outcome than OS twins for nongenetic reasons, most obviously via any consequence of being the same sex. We addressed this challenge in two ways. First, we standardized test scores within sexes, so twin similarity would be measured independently from any mean or variance differences between sexes. Second, differences between SS and OS twins were compared with differences between SS and OS siblings who are close in age but not twins.

The size and representativeness of administrative data are benefits to be weighed against the limitation of not having zygosity measures. That said, studying the Scarr-Rowe interaction with data in which zygosity is known involves other assumptions, which may be more problematic than is broadly appreciated. The variance components estimated in this work are population parameters, as are terms for the moderation of these components by SES. However, the design and differential recruitment of many twin studies complicate the definition of the target population that the estimates produced by these studies are supposed to represent. For that matter, given the strong association between SES and test scores, estimated interactions have some dependence on the metrics by which SES and test scores are specified in models, but there has been little articulated reason beyond convention to favor any particular metric over others. None of this is intended to dismiss previous work; however, drawing inferences about biosocial interaction from models estimating statistical interactions is a thorny matter, involving substantial assumptions that are not simply dispelled by having data on twins' zygosity. Instead, evidence across multiple research designs is needed (22, 23).

Results

Fig. 1 shows the relationship between one measure of SES, maternal education, and scores on the Florida Comprehensive Assessment Test (FCAT) for all pairs in the sample [SI Appendix contains parallel results for the SES measure based on median income, principal components analysis (PCA)-based SES indexes, and for an alternative achievement test]. Many studies show positive relationships between these various SES measures and cognitive functioning (24–29). Children whose mothers did not finish high school were ~ 0.5 SDs below the overall mean, compared with ~ 0.7 SDs above the mean for children whose mothers completed college.

Pair-level ICCs in Fig. 2 show that data are consistent, with substantial genetic influence on test scores. For twins, a sizable difference exists between the ICCs for SS compared with OS twins. The difference is several times larger than the corresponding difference between SS and OS nontwins, as we would expect if the primary cause of the twin difference is that approximately half of the SS twins are MZ twins. Nevertheless, we can also see that SS nontwins are more similar than OS nontwins and that OS twins are more similar than OS nontwins. This pattern implies that sex similarity and twin status have their own, albeit small, influence on pair similarity.

Data are thus consistent, with familiar results that cognitive achievement differs over SES and that MZ twins have more similar cognitive achievement than DZ twins. To consider whether the data also provide evidence of a Scarr-Rowe interaction, we followed Turkheimer and Horn's review of the literature, which summarizes existing evidence as a combination of two phenomena (21). We consider each in turn, as shown in Fig. 3, which presents within- and between-pair variances for SS and OS twins.

First, Turkheimer and Horn indicate that “the between-pair variance of MZ pairs decreases in poor environments” (ref. 21, p. 63). Contrary to this relationship, we found that the between-pair variance of SS twins is actually lowest in the highest SES families. Given that SS twins are a relatively equal combination of MZ and DZ twins, one possibility is that a pattern supporting the hypothesis among MZ SS twins is masked by an even stronger pattern in the opposite direction among DZ SS twins. However, Fig. 3 shows that corresponding results for OS twins (all of whom are DZ) give no indication of such a pattern. Between-pair variances in achievement test scores for high-school educated parents of OS twins are higher in all cases than it is for parents without a high school diploma.

Second, Turkheimer and Horn report that “the within-pair variance of MZ twin pairs increases at lower levels of SES: poverty appears to have the effect of making MZ twins more different from each other” (ref. 21, p. 61). We would therefore expect in our data that the within-pair variance for SS twins whose mother did not graduate from high school would be higher than the variance for SS twins whose mother has a high school diploma. However, this is not the case in any of the SS twin comparisons shown in Fig. 3.

As before, one might be concerned that a contrary pattern among DZ twins is masking the effect, because some SS twins are DZ. However, this possibility is again contradicted by results shown in Fig. 3 for OS twins, in which the pattern of results is opposite of what would be necessary for such masking to happen. We found very similar results for alternative measures of SES (SI Appendix, Figs. S10, S18, and S22).

For older children and math tests, there may be an indication of higher within-pair variance for SS twins with high school-versus college-educated mothers. This is the result closest to the expectations of the hypothesis. Even here, however, we observed that a similar pattern was present in these cases for OS twins, which strengthens the alternative possibility that the result among SS twins could be driven by DZ pairs, rather than by MZ pairs as entailed by a Scarr-Rowe interaction.

Turkheimer and Horn (21) showed that the combined result of changing variance is an increasing divergence of ICCs between

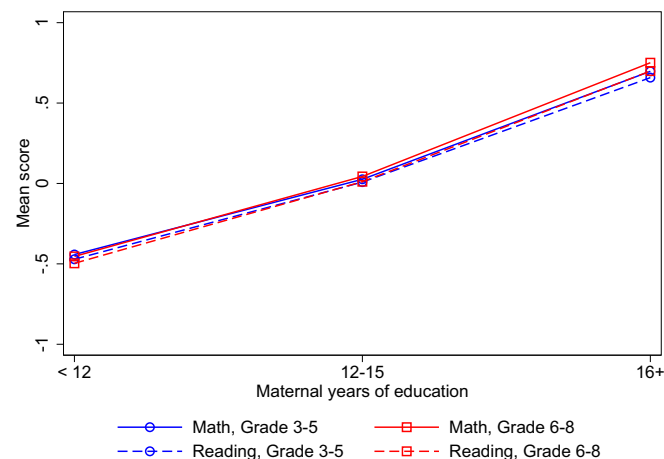


Fig. 1. Maternal years of education and average achievement test score for combined twin and sibling pairs sample. The means of gender-standardized test scores in mathematics and reading over maternal years of education (separated by age group and test type) are shown. The sample includes all twin pairs and closely spaced sibling pairs with available test scores. A closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. $n = 299,426$ children (24,640 twins and 274,786 singletons) and 1,796,532 children-year observations.

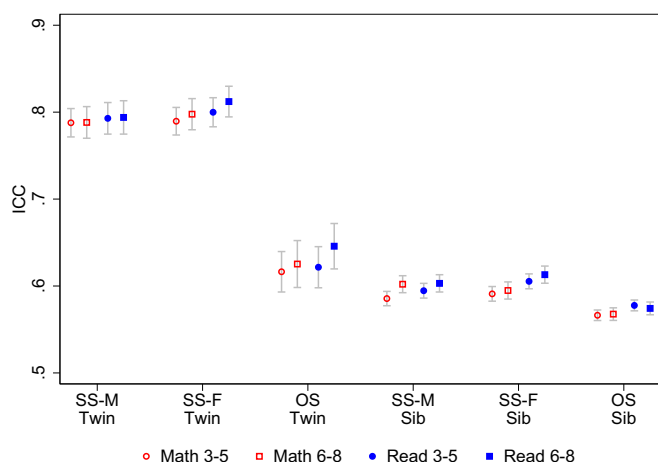


Fig. 2. ICCs and 95% CIs in relation to math and reading achievement tests for SS-M, SS-F, and OS samples of younger (grades 3–5) and older (grades 6–8) twins and closely spaced siblings. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. ICCs are based on multilevel mixed-effects linear regression estimated with maximum likelihood, where within-individual across-grades errors are assumed to have an autoregressive structure of order one. Random effects are structured at the twin/sibling pair and individual levels. ICC is computed as the ratio of between-pair variation to the sum of within- and between-pair variations. $n = 299,426$ children (24,640 twins and 274,786 singletons) and 1,796,532 children-year observations. F, female; M, male; Read, reading; Sib, sibling.

MZ and DZ twins, because such divergence implies higher heritability estimates. Accordingly, we would expect the difference in ICCs between SS and OS twins to diverge as SES increases. We present ICCs by maternal education in Fig. 4, and contrary to expectations from Turkheimer and Horn, we observed the opposite pattern. Fig. 4 also shows that differences are

much smaller and inconsistently signed between SS and OS nontwin siblings, which contradicts the possibility that our results diverge from the hypothesis due to SES moderation of sex difference per se. We found very similar results for alternative measures of SES (*SI Appendix*, Figs. S11, S19, and S23).

A different method of estimating the interaction is to adapt a model presented by Purcell (30) and used in the meta-analysis by Tucker-Drob and Bates (16). This model extends the conventional ACE model that estimates additive genetic effects (A), shared environment (C), and nonshared environment (E). The extended model includes a separate term for variance accounted for by the SES measure (M), along with terms for moderation by SES (A', C', E'). In these models, "A" indicates the estimated narrow-sense heritability for someone with average SES. This method is more conventional but also involves stronger metric assumptions. These results are presented in Table 1.

The prediction is that the moderator term for the additive genetic component (A') would be positively signed and statistically significant. Our results over different outcomes uniformly contradict this prediction. We also tested whether our results depend on assumptions about the proportion of MZ twins in the SS twin group; however, over the range of plausible assumed values we tested (0.4–0.6), the interaction terms were uniformly signed in the wrong direction from the prediction, but not always statistically significantly so (*SI Appendix*, Table S3). We also explored two alternative, more continuous measures of SES based on PCA analyses. Again, the interaction terms were uniformly signed in the wrong direction and not always significant (*SI Appendix*, Tables S4–S7).

To address the possibility that results may be confounded in one way or another by SES differences by race, we also conducted analyses in which we restricted the sample to children with white mothers (whites are the largest race/ethnic group in the sample). We further conducted analyses restricting the sample to mothers up to age 30 y, for whom rates of in vitro fertilization (IVF) use are extremely low (31), in case SES differences in the use of IVF treatments could distort results for

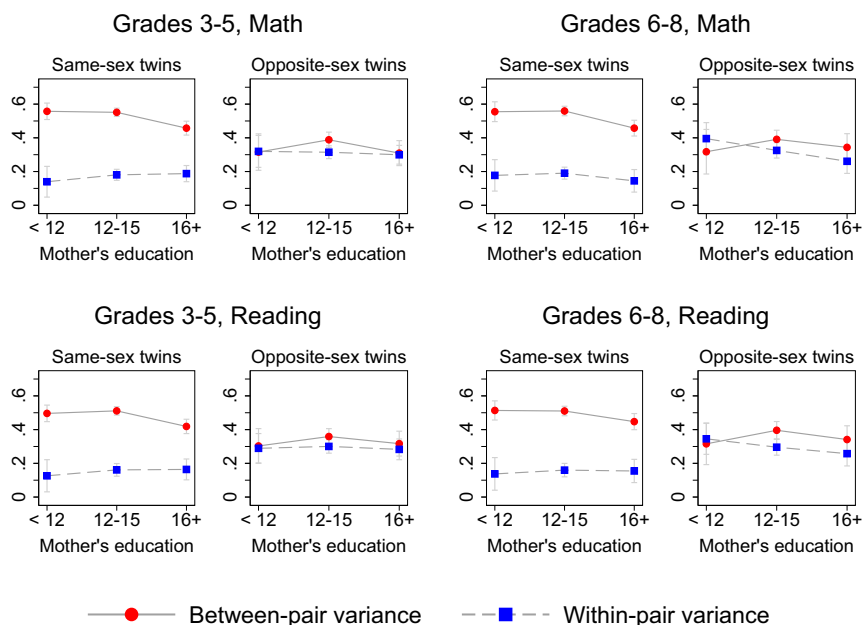


Fig. 3. Estimates and 95% CIs of between- and within-pair variations for SS and OS twin pairs and achievement test scores in mathematics and reading assessed in grades 3–5 or 6–8 and split by years of maternal education. The variances were obtained using multilevel mixed-effects linear regression estimated with maximum likelihood, where within-individual across-grades errors are assumed to have an autoregressive structure of order one. Random effects are structured at the twin pair and individual levels. $n = 24,640$ twins and 147,828 children-year observations.

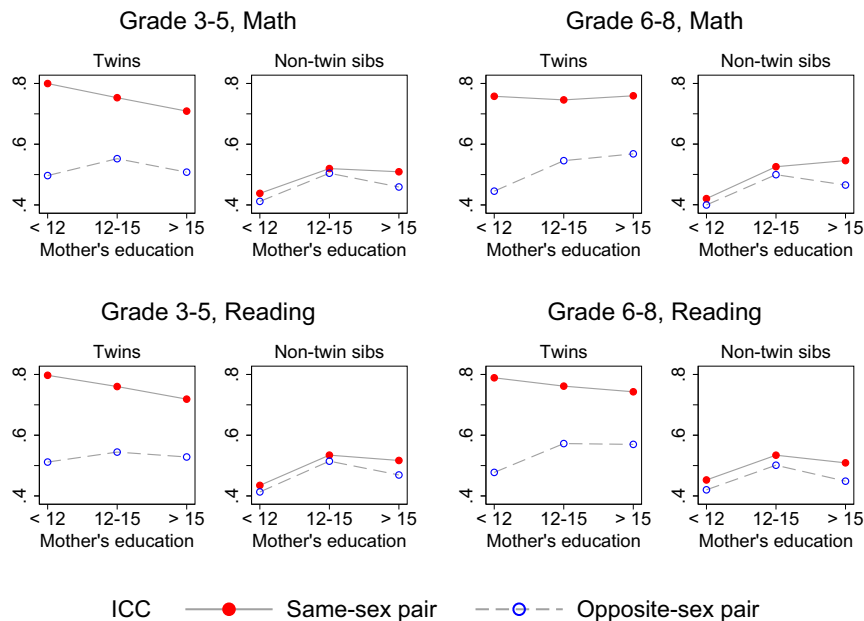


Fig. 4. Estimates of ICCs for SS and OS pairs of twins and nontwin siblings based on test scores in mathematics and reading assessed in grades 3–5 or 6–8 and split by years of maternal education. Siblings are defined as two individuals having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. The ICCs are based on multilevel mixed-effects linear regression estimated with maximum likelihood, where within-individual across-grades errors are assumed to have an autoregressive structure of order one. Random effects are structured at the twin/sibling pair and individual levels. ICC is computed as the ratio of between-pair variation to the sum of within- and between-pair variations. $n = 299,426$ children (24,640 twins and 274,786 singletons) and 1,796,532 children-year observations.

DZ twins (DZ, but not MZ, twinning rates are much higher when IVF is used). In both cases, as well as in other analyses also reported in the [SI Appendix](#), our substantive conclusion remains the same: no evidence of a Scarr-Rowe interaction.

Discussion

Although past research has found stronger evidence for a Scarr-Rowe interaction in the United States than elsewhere, we failed to find evidence of increasing genetic influence on the cognitive similarity of twins as SES increases. Our analysis makes use of large-scale population-level administrative data in the United States where population inferences are not compromised by patterns of successful recruitment into a survey, especially among low-SES families. Our results suggest that the mixed results in this literature cannot be explained by lack of SES diversity in some samples or by differences between the United States and other nations.

Trying to understand why we failed to find an interaction when (some) others have found an interaction is difficult. For example,

Kovas et al. (32) reported higher grade-school heritabilities for literacy and numeracy versus IQ test scores, raising the possibility that our results could differ if we had IQ test results instead of math and reading achievement tests. However, the differences reported in that paper are present for younger but not older children, whereas our results are consistent across ages 8–14 y. Reviews of the literature have also cited other studies using achievement tests as evidence for the interaction (21).

Perhaps our lack of direct zygosity measurement is the culprit, but the study that launched this literature also compared SS and OS twins (1), and our study's sample size surmounts the concerns about statistical power raised regarding that study (33). Maybe the absence of private schools in our sample is an issue, because children of higher-SES families are more likely to attend private schools. However, decades of research have failed to establish clear evidence of a broadly positive causal effect of private compared with public schools in the United States, especially for children of more affluent families (34, 35).

Table 1. Modified ACE variance components model

Test	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (nonshared environment)	A' (genetic \times SES)	95% CI for A'		n
							Lower	Upper	
Math	3–5	0.590	0.143	0.052	0.209	−0.101	−0.120	−0.082	34,432
Reading	3–5	0.619	0.149	0.003	0.224	−0.030	−0.073	0.013	
Math	6–8	0.628	0.169	0.000	0.181	−0.050	−0.110	0.010	21,653
Reading	6–8	0.594	0.166	0.017	0.197	−0.077	−0.130	−0.023	

This table is based on variance components models used in Purcell (30) and Tucker-Drob and Bates (16). The first two models (grades 3–5) use a relatedness value of 0.76, whereas the latter two models (grades 6–8) use a relatedness value of 0.77. Relatedness is defined as $(n - M)/n + 0.5(M/n)$, where n is the number of SS twins and M is the number of OS twins. The first two columns list test and test grades. Subsequent columns present variances due to additive genetic effects (A); SES measure, which in our case is years of maternal education at birth (M); shared environment (C); and nonshared environment (E). A' is a moderator term of interest for additive genetic effects displayed together with lower and upper bounds of 95% CIs in the following two columns. The Scarr-Rowe hypothesis requires A' to be positive and statistically significant. The last column presents sample sizes used in estimation.

At a minimum, our findings indicate that the nature of the gene-environment interaction is less clear cut than may have been supposed from smaller samples. How to effectively describe the interplay of genes and environments remains a profound scientific challenge, and we hope continued improvements in the availability of both administrative and genomic data will yield important progress in the years ahead. One potential frontier for future work would make use of molecular genetic data rather than studies of twin similarity. Although early efforts to identify genetic correlates of cognitive ability faced substantial challenges (36), current efforts are quite promising (37–39). As causal genetic variants for cognition become better established, one question will be whether they are more influential in some environments than others and whether there will be a more systematic genomic pattern of greater cumulative influence in higher-SES environments. However, it may still be some time before sufficient sample sizes will be available to investigate this question fully.

Materials and Methods

Data for the study are based on confidential matched birth and school records obtained for all children who were born in Florida between 1994 and 2002 and subsequently educated in a Florida public school. Birth certificates were matched to the Education Data Warehouse by the Florida Department of Health and the Florida Department of Education using four variables: first and last name, date of birth, and social security number. To maximize correct matches, transposition of letters or numbers was allowed in up to two instances as long as the transposition did not match more than one record. Children were included if they (i) were born in Florida, (ii) remained in Florida until school age, and (iii) attended Florida public schools. This study was approved by the institutional review boards of Northwestern University, the University of Florida, and Florida's education and health agencies and involved a thorough review of the Family Educational Rights and Privacy Act, as well as the Health Insurance Portability and Accountability Act regulations.

Twins were ascertained by plurality information on the birth certificate, as well as the same date of birth and maternal characteristics. Nontwin siblings were ascertained by residential address in schooling and were linked back to the relevant students' birth records to check that students we believed to be siblings were actually siblings (e.g., by comparing maternal characteristics such as date of birth). For families with more than two nontwin siblings in eligible cohorts, we constructed our sample of nontwin-sibling pairs by taking the two who were closest in age.

Overall, 80.7% of all children born in Florida, and 79.5% of all twins born in Florida, were matched to school records. This aligns closely with the 80.9% rate of kindergarten-age children born in Florida and attending Florida public schools based on data from the American Community Survey.

Florida has a larger population (20.3 million in 2015) than several nations in which the Scarr-Rowe interaction has been previously studied (e.g., The Netherlands, Sweden). Relative to other US states, Florida has higher than national average economic inequality, child poverty rates, and percentage of nonwhite residents, making it especially promising for assessing potential socioeconomic heterogeneity in causal effects. The sample's diversity on key correlates of SES such as race, mother's education, and mother's age is shown in *SI Appendix, Table S1*. The table also shows that characteristics associated with disadvantage are slightly overrepresented in our estimation sample in comparison with all Florida births, which predominantly reflects the fact that families with adverse characteristics are more likely to enroll children in public school and less likely to migrate out of Florida.

The measures of SES used in the paper are derived from birth records: the mother's educational attainment, the median income of the zip code of the

mother's residence, and two PCA-based composite measures of multiple SES inputs. Maternal education is grouped into three categories: high school dropout (<12 y of education), high school graduate (12–15 y of education), and college graduate (≥ 16 y of education). Neighborhood income was determined by using the median income value based on the 2000 US Census for each zip code of parental primary residence as indicated in birth records. The primary composite measure includes these two measures along with family structure [mother and father married, parents unmarried but father listed on birth certificate, no father listed (reference category)] and whether the birth was paid for by Medicaid. Maternal education in years was used. Because zip code income is measured at different levels (zip code) than other SES inputs (individual) we also explored PCA analysis excluding median income.

The primary academic achievement variables used were the reading and mathematics scores from the FCAT, which was administered annually throughout Florida public schools over our study period. Published correlations between tests designed to measure general cognitive ability (i.e., "IQ", g) and standardized achievement test scores vary but are often ~ 0.7 (40, 41), with lower correlations observed when short or limited-domain IQ tests are used or when academic achievement is measured from grade-point average or teacher assessment rather than testing.

We standardized results for both FCAT tests by grade, year, and sex, to have a mean of zero and an SD of one in the overall population. This meant that the mean score in our linked sample would be above zero, because children of migrants into Florida have lower achievement than children born in Florida. We averaged scores across multiple grades to reduce measurement error. Stanford achievement tests were also administered to Florida students in school years 2000/2001–2006/2007, and we include results from these in *SI Appendix, Figs. S4–S9, S12, S13, S20, S21, S24, and S25*. Because of the restriction on the years when the test was administered, we were only able to perform the analysis for grades 3–5.

We used two estimation strategies which are described in more detail in *SI Appendix, Technical Appendix*. First, we divided pairs into SES subgroups and we computed between-pair and within-pair variances for each group using mixed-effects linear regression estimated with maximum likelihood where within-individual across-grades errors are assumed to have an autoregressive structure of order one. Random effects are structured at the twin/sibling pair and individual levels. We also reported the ICCs, which are computed as the ratio of between-pair variation to the sum of within- and between-pair variations.

Second, we used a model of continuous moderation, as described by Tucker-Drob and Bates (16). This model of phenotype variation starts with the A (additive genetic), C (shared environment), E (nonshared environment) parameters of the classic twin study model and adds a parameter M to separate the variance accounted for by the potential moderating variable (in our case, the SES measure). Then it adds interaction parameters A' , C' , E' , so that, for example, $A' = A \times \text{SES}$. Identifying these parameters requires assuming the average relatedness of SS twin pairs, which are a combination of MZ ($r = 1$) and DZ ($r = 0.5$) twins. In the main analysis, we used relatedness of 0.76 (grades 3–5) and 0.77 (grades 6–8) derived on the basis of our data but also conducted sensitivity estimates over a range of values between 0.700 and 0.800.

ACKNOWLEDGMENTS. We thank Timothy Bates, Elliot Tucker-Drob, and Eric Turkheimer for helpful discussions; Livia Baer-Bositis for editorial assistance; and Elliot Tucker-Drob for additional help with MPlus code. We thank the Florida Departments of Education and Health for providing the de-identified, matched data used in this analysis. D.N.F. and J.R. acknowledge support from the National Science Foundation and the Institute for Education Sciences (CALDER Grant), and D.N.F. acknowledges support from the National Institute of Child Health and Human Development and the Bill and Melinda Gates Foundation. The conclusions expressed in this paper are those of the authors and do not represent the positions of the Florida Departments of Education and Health or those of our funders.

- Scarr-Salapatek S (1971) Race, social class, and IQ. *Science* 174:1285–1295.
- Guo G, Stearns E (2002) The social influences on the realization of genetic potential for intellectual development. *Soc Forces* 80:881–910.
- Nisbett RE, et al. (2012) Intelligence: New findings and theoretical developments. *Am Psychol* 67:130–159.
- Cunha F, Heckman J (2007) The technology of skill formation. *Am Econ Rev* 97:31–47.
- Turkheimer E, Haley A, Waldron M, D'Onofrio B, Gottesman II (2003) Socioeconomic status modifies heritability of IQ in young children. *Psychol Sci* 14:623–628.
- Bates TC, Lewis GJ, Weiss A (2013) Childhood socioeconomic status amplifies genetic effects on adult intelligence. *Psychol Sci* 24:2111–2116.
- Friend A, DeFries JC, Olson RK (2008) Parental education moderates genetic influences on reading disability. *Psychol Sci* 19:1124–1130.

- Harden KP, Turkheimer E, Loehlin JC (2007) Genotype by environment interaction in adolescents' cognitive aptitude. *Behav Genet* 37:273–283.
- Rowe DC, Jacobson KC, Van den Oord EJ (1999) Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. *Child Dev* 70:1151–1162.
- Tucker-Drob EM, Rhemtulla M, Harden KP, Turkheimer E, Fask D (2011) Emergence of a gene \times socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychol Sci* 22:125–133.
- Grant MD, et al. (2010) Does parental education have a moderating effect on the genetic and environmental influences of general cognitive ability in early adulthood? *Behav Genet* 40:438–446.
- Hanscombe KB, et al. (2012) Socioeconomic status (SES) and children's intelligence (IQ): In a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS One* 7:e30320.

13. Rhemtulla M, Tucker-Drob EM (2012) Gene-by-socioeconomic status interaction on school readiness. *Behav Genet* 42:549–558.
14. van der Sluis S, Willemsen G, de Geus EJ, Boomsma DI, Posthuma D (2008) Gene-environment interaction in adults' IQ scores: Measures of past and present environment. *Behav Genet* 38:348–360.
15. Kirkpatrick RM, McGue M, Iacono WG (2015) Replication of a gene-environment interaction via multimodel inference: Additive-genetic variance in adolescents' general cognitive ability increases with family-of-origin socioeconomic status. *Behav Genet* 45:200–214.
16. Tucker-Drob EM, Bates TC (2016) Large cross-national differences in gene x socioeconomic status interaction on intelligence. *Psychol Sci* 27:138–149.
17. Tucker-Drob EM, Briley DA, Harden KP (2013) Genetic and environmental influences on cognition across development and context. *Curr Dir Psychol Sci* 22:349–355.
18. Chapple S, Richardson D (2009) *Doing Better for Children* (OECD Publishing, Paris).
19. Figlio D, Guryan J, Karbownik K, Roth J (2014) The effects of poor neonatal health on children's cognitive development. *Am Econ Rev* 104:3921–3955.
20. Hart SA, Soden B, Johnson W, Schatschneider C, Taylor J (2013) Expanding the environment: Gene \times school-level SES interaction on reading comprehension. *J Child Psychol Psychiatry* 54:1047–1055.
21. Turkheimer E, Horn E (2014) Interactions between socioeconomic status and components of variation in cognitive ability. *Behavior Genetics of Cognition Across the Lifespan*, ed Reynolds C (Springer, New York), pp 41–68.
22. Benjamin B, Deary IJ, Visscher PM (2006) Precision and bias of a normal finite mixture distribution model to analyze twin data when zygosity is unknown: Simulations and application to IQ phenotypes on a large sample of twin pairs. *Behav Genet* 36:935–946.
23. Benjamin B, Wilson V, Whalley LJ, Visscher PM, Deary IJ (2005) Large, consistent estimates of the heritability of cognitive ability in two entire populations of 11-year-old twins from Scottish mental surveys of 1932 and 1947. *Behav Genet* 35:525–534.
24. Duncan GJ, Magnuson K (2012) Socioeconomic status and cognitive functioning: Moving from correlation to causation. *Wiley Interdiscip Rev Cogn Sci* 3:377–386.
25. Hackman DA, Farah MJ, Meaney MJ (2010) Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nat Rev Neurosci* 11:651–659.
26. Larson K, Russ SA, Nelson BB, Olson LM, Halfon N (2015) Cognitive ability at kindergarten entry and socioeconomic status. *Pediatrics* 135:e440–e448.
27. Mani A, Mullainathan S, Shafir E, Zhao J (2013) Poverty impedes cognitive function. *Science* 341:976–980.
28. Noble KG, Houston SM, Kan E, Sowell ER (2012) Neural correlates of socioeconomic status in the developing human brain. *Dev Sci* 15:516–527.
29. Ruiz JdelC, Quackenboss JJ, Tulve NS (2016) Contributions of a child's built, natural and social environments to their general cognitive ability: A systematic scoping review. *PLoS One* 11:e0147741.
30. Purcell S (2002) Variance components models for gene-environment interaction in twin analysis. *Twin Res* 5:554–571.
31. Bitler MP, Schmidt L (2012) Utilization of infertility treatments: The effects of insurance mandates. *Demography* 49:125–149.
32. Kovas Y, et al. (2013) Literacy and numeracy are more heritable than intelligence in primary school. *Psychol Sci* 24:2048–2056.
33. Eaves LJ, Jinks JL (1972) Insignificance of evidence for differences in heritability of IQ between races and social classes. *Nature* 240:84–88.
34. Eppl D, Romano R, Zimmer R (2016) Charter schools: A survey of research on their characteristics and effectiveness. *Handbook of the Economics of Education*, eds Hanushek E, Machin S, Woessmann L (Elsevier, Amsterdam), pp 139–208.
35. Eppl D, Romano R, Urquiola M (2017) School vouchers: A survey of the economics literature. *J Econ Lit* 55:441–492.
36. Chabris CF, et al. (2012) Most reported genetic associations with general intelligence are probably false positives. *Psychol Sci* 23:1314–1323.
37. Davies G, et al. (2015) Genetic contributions to variation in general cognitive function: A meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Mol Psychiatry* 20:183–192.
38. Selzam S, et al. (2017) Predicting educational achievement from DNA. *Mol Psychiatry* 22:267–272.
39. Rietveld CA, et al. (2014) Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc Natl Acad Sci USA* 111:13790–13794.
40. Naglieri J, Bronstein B (2003) Intelligence and achievement: Just how correlated are they? *J Psychoeduc Assess* 21:244–260.
41. Kaufman S, Reynolds M, Liu X, Kaufman A, McGrew K (2002) Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence* 40:123–138.

Supporting Information:

Technical Appendix

Intraclass correlations

The intraclass correlations presented in Figure 2 are based on multilevel mixed-effects linear regressions fitted using maximum likelihood. The models include three levels: multiple years of test scores (either grades 3 to 5 or grades 6 to 8) are nested within individuals, and individuals are nested within twin or sibling pairs. Accordingly, the analysis includes random intercepts for both an individual and a pair. The regression takes the following form:

$$Y_{gij} = \alpha_{000} + \mu_{00j} + \pi_{0ij} + \varepsilon_{gij} \quad (1)$$

with subscripts g used for grade, i for individual, and j for pair. In the regression, α_{000} is a fixed intercept representing grand mean; μ_{00j} is pair specific random intercept (pair deviation from fixed intercept); π_{0ij} is an individual specific random intercept (individual deviation from pair predicted outcome); ε_{gij} is a residual (grade specific deviation from individual predicted outcome). Level 1 of the model is $Y_{gij} = \beta_{0ij} + \varepsilon_{gij}$; level 2 is $\beta_{0ij} = \delta_{00j} + \pi_{0ij}$; and level 3 is $\delta_{00j} = \alpha_{000} + \mu_{00j}$. We model the residuals to have autoregressive structure of order one. We fit separately $2 \times 2 \times 2 \times 3$ combinations:

- Two outcomes: reading and math
- Two grade levels: 3 to 5 or 6 to 8
- Two pair types: twins and non-twin siblings
- Three gender-composition types: male-male, female-female, male-female

This model allows us to compute within-person, within-pair and between pair variances for each of the 24 combinations listed above. Namely, level 1 variance (within person) is V_e ; level 2 variance (within pair) is $V_{\pi 0}$; and level 3 variance (between pair) is $V_{\mu 00}$. Once we obtain the three variances from the model we can compute total pair variance $V_{total} = V_{\pi 0} + V_{\mu 00}$, which is just a sum of within and between pair variances. The intraclass correlation is then a ratio of between pair variance to total pair variance i.e. $ICC = \frac{V_{\mu 00}}{V_{total}}$.

Variance decomposition

We use the same method and exactly the same equation for the variance decomposition shown in Figures 3 and 4, with one specific modification. Namely, we compute variances for $2 \times 2 \times 2 \times 2 \times 3$ groups:

- Two outcomes: reading and math
- Two grade levels: 3 to 5 and 6 to 8

- Two gender composition types: same-sex and opposite-sex
- Two pair types: twins and non-twin siblings
- Three levels of SES: low, medium, high

For the SES levels we repeat the analysis separately for maternal education (high school dropout, high school graduate, college graduate), zip code income terciles, and PCA'ed SES index terciles.

In Figure 3 we directly plot between and within variances obtained from the model for the twin sample, which yields $3 \times 2 \times 2 \times 2 = 48$ estimates.

In Figure 4 we compute ICCs for each SES level, pair type, outcome, grade level and sex composition, which again yields 48 estimates. ICCs are computed as a ratio of between pair variance to total pair variance for each of the groups.

We compute ICCs and perform variance decomposition using Stata, and the code is available online.

ACE model with continuous moderator

The standard ACE model decomposes genetic phenotype into three factors: additive genetic factors (A), common environment factors (C), and specific environmental factors (E). It can be written mathematically as $V_p = V_A + V_C + V_E$, where V_p is a total variance of the trait and the three linearly additive variances represent factors A , C and E . This model implies that $V_A + V_C + V_E = 1$, and heritability is defined as $H = \frac{V_A}{V_p}$.

In our particular case, we are seeking to understand whether genetic influence (A) is moderated by socioeconomic status of family (SES moderator M). Since we are interested in the interactions with childhood conditions, we expand the model to include A' , C' and E' components where $A' = A \times M$; $C' = C \times M$; $E' = E \times M$. Central to our interest is an interaction term A' that represents gene \times SES interaction. We use two continuous SES moderators in this analysis: maternal years of education and PCA'ed SES index. Thus, the variance in expanded model changes to $V_{p|M} = (A + A')^2 + (C + C')^2 + (E + E')^2$, and the standardization condition takes a form of $\frac{(A+A')^2}{V_{p|M}} + \frac{(C+C')^2}{V_{p|M}} + \frac{(E+E')^2}{V_{p|M}} = 1$.

Identification of parameters for this model with twin data usually has used information on zygosity of twins, relying on the fact that monozygotic twins (MZ) share all their genes, while dizygotic twins (DZ) share on average half of their genes by descent. Thus, for MZ twins the observed covariance between twins equals $V_A + V_C$ while for DZ twins this covariance is $0.5 \cdot V_A + V_C$. Similarly when conditioning on continuous moderator we simply replace V_A and V_C with $(A + A')^2$ and $(C + C')^2$, respectively.

In the absence of information on zygosity, we use the observed gender composition of twin pair as a proxy to fit the model whose results are shown in Table 1. This is possible because all opposite-sex twin pairs are dizygotic (DZ), while same-sex twin pairs are a mix of dizygotic (DZ) and monozygotic (MZ) phenotypes. We standardize test scores within sex to remove any mean differences between sexes.

In order to proceed with modeling, however, we need a value for rSS which is genetic relatedness of the same-sex twin group. It can be estimated from $rSS = \frac{N-M}{N} + 0.5\frac{M}{N}$, where N is the number of same-sex twins in the sample and M is the number of opposite-sex twins in the sample. The first right-hand side term is a proportion of monozygotic (MZ) twins in the group, while the latter is a proportion of dizygotic (DZ) twins (multiplied by their relatedness of .5). Given that same-sex twins are often composed of about equal numbers of MZ and DZ twins, .75 is sometimes used as an approximation of rSS . We compute rSS directly from our data, and in the main sample that yields a value of 0.76 in grades 3 to 5 and value of 0.77 in grades 6 to 8. We test the robustness of the results for alternative values of rSS : 0.700, 0.725, 0.750, 0.775, 0.800. Based on this we run $2 \times 2 \times 6$ models using twin data:

- Two outcomes: reading and math
- Two grade levels: 3 to 5 and 6 to 8
- Six values for rSS assumption

We compute these models using MPlus, and the code is available online.

Figures and tables

Table S1: Demographic characteristics of mothers

	(1)	(2)	(3)	(4)
	Births 1994-2002			
Characteristic	All	Matched to school records	Final twin sample	Final sibling sample
African-American	21.9	24.1	25.6	30.7
Hispanic	23.8	24.2	18.0	25.2
Immigrant	24.0	23.4	18.0	22.2
Married	65.2	62.4	68.0	62.2
HS dropout	20.1	21.8	15.2	25.4
College graduate	21.3	18.0	24.1	18.4
Age 21 or below	21.4	23.2	14.5	24.1
Age 37 or above	7.5	7.1	10.7	4.8
N	1,636,968	1,312,345	24,640	274,786

Notes: The first column presents fractions in total population of children born in Florida between 1994 and 2002. The second column presents fractions in total population of children born between 1994 and 2002 linked to Florida school records. The third column presents fractions in final twin sample used in the empirical analysis. The fourth column presents fractions in final sibling sample used in the analysis.

Table S2: Alternative variance components model. Moderator: Maternal years of education

Test	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-shared environment)	A' (genetic x SES)	95% CI for A'		C' (common env. x SES)	95% CI for C'		E' (non-shared env. x SES)	95% CI for E'		N
							Lower	Upper		Lower	Upper		Lower	Upper	
Math	3 to 5	0.590	0.143	0.052	0.209	-0.101	-0.120	-0.082	0.118	0.086	0.150	0.030	0.020	0.040	34,432
Reading	3 to 5	0.619	0.149	0.003	0.224	-0.030	-0.073	0.013	0.151	0.126	0.175	0.016	-0.005	0.037	
Math	6 to 8	0.628	0.169	0.000	0.181	-0.050	-0.110	0.010	-0.173	-0.206	-0.139	-0.018	-0.048	0.012	21,653
Reading	6 to 8	0.594	0.166	0.017	0.197	-0.077	-0.130	-0.023	0.166	0.119	0.214	0.036	0.012	0.060	

Notes: This table is based on variance components model. Column (1) lists a test - either mathematics or reading while column (2) lists test grades. The first two models (grades 3 to 5) use relatedness value of 0.76 while the latter two models (grades 6 to 8) use relatedness value of 0.77. Relatedness is defined as $R_{ss} = \frac{N-M}{N} + 0.5\frac{M}{N}$ where N is number of same-sex twins and M is number of opposite-sex twins. Subsequent columns present variances due to additive genetic effects (A ; column 3); socioeconomic status measure which in this case is years of maternal education at birth (M ; column 4); variances due to shared environment (C ; column 5) and non-shared environment (E ; column 6). Columns (7) to (9) present a moderator term of interest for additive genetic effects (A') with 95% confidence interval. Columns (10) to (12) present a moderator term for shared environment (C') with 95% confidence interval. Columns (13) to (15) present moderator term for non-shared environment (E') with 95% confidence interval. Column (16) presents sample sizes used in estimation. Scarr-Rowe hypothesis requires estimate in column (7) to be positive and statistically significant.

Table S3: Robustness of alternative variance components model. Moderator: Maternal years of education

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Test	R _{SS}	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-sh. environment)	A' (genetic x SES)	95% CI for A'		C' (common env. X SES)	95% CI for C'		E' (non-sh. env. X SES)	95% CI for E'		N
Math	0.700	3 to 5	0.706	0.144	0.004	0.142	-0.069	-0.094	-0.044	0.141	0.113	0.170	0.030	0.010	0.049	34,432
Reading		3 to 5	0.674	0.148	0.001	0.174	-0.037	-0.061	-0.012	0.143	0.118	0.169	0.026	0.008	0.044	
Math		6 to 8	0.691	0.168	0.000	0.123	-0.052	-0.083	-0.020	-0.164	-0.195	-0.132	-0.012	-0.038	0.015	21,653
Reading		6 to 8	0.674	0.166	0.001	0.136	-0.049	-0.082	-0.016	0.173	0.144	0.202	0.037	0.010	0.064	
Math	0.725	3 to 5	0.669	0.144	0.014	0.167	-0.081	-0.107	-0.054	0.137	0.104	0.169	0.030	0.014	0.047	34,432
Reading		3 to 5	0.651	0.148	0.001	0.195	-0.036	-0.065	-0.007	0.146	0.121	0.171	0.022	0.004	0.041	
Math		6 to 8	0.667	0.168	0.000	0.144	-0.051	-0.087	-0.015	-0.167	-0.198	-0.136	-0.014	-0.040	0.011	21,653
Reading		6 to 8	0.651	0.166	0.001	0.158	-0.052	-0.092	-0.012	0.175	0.145	0.205	0.033	0.006	0.060	
Math	0.750	3 to 5	0.613	0.143	0.040	0.197	-0.096	-0.117	-0.075	0.123	0.090	0.157	0.031	0.020	0.042	34,432
Reading		3 to 5	0.629	0.148	0.002	0.215	-0.033	-0.071	0.005	0.149	0.125	0.174	0.018	-0.002	0.038	
Math		6 to 8	0.645	0.168	0.000	0.166	-0.050	-0.095	-0.006	-0.170	-0.202	-0.138	-0.017	-0.043	0.010	21,653
Reading		6 to 8	0.626	0.166	0.004	0.179	-0.059	-0.111	-0.008	0.175	0.139	0.210	0.032	0.004	0.060	
Math	0.775	3 to 5	0.557	0.143	0.070	0.226	-0.106	-0.123	-0.089	0.110	0.080	0.141	0.028	0.019	0.037	34,432
Read		3 to 5	0.601	0.149	0.008	0.237	-0.020	-0.067	0.026	0.153	0.129	0.177	0.011	-0.010	0.032	
Math		6 to 8	0.623	0.169	0.000	0.185	-0.049	-0.116	0.018	-0.174	-0.208	-0.139	-0.019	-0.051	0.013	21,653
Reading		6 to 8	0.581	0.165	0.024	0.203	-0.083	-0.131	-0.035	0.162	0.113	0.210	0.037	0.017	0.058	
Math	0.800	3 to 5	0.507	0.142	0.096	0.250	-0.112	-0.128	-0.096	0.101	0.073	0.129	0.024	0.016	0.032	34,432
Reading		3 to 5	0.521	0.147	0.061	0.267	-0.091	-0.109	-0.073	0.105	0.069	0.140	0.033	0.025	0.041	
Math		6 to 8	0.593	0.169	0.008	0.206	-0.027	-0.122	0.068	-0.179	-0.209	-0.150	-0.029	-0.066	0.008	21,653
Reading		6 to 8	0.518	0.164	0.063	0.230	-0.103	-0.130	-0.076	0.141	0.101	0.181	0.038	0.027	0.049	

Notes: This table is based on variance components model. Column (1) lists a test - either mathematics or reading while column (3) lists test grades. Column (2) presents assumptions regarding relatedness of same-sex twin group. For each combination of test type and grade range we test robustness to five relatedness values. The preferred relatedness values in Table 1 are 0.76 (grades 3 to 5) and 0.77 (grades 6 to 8). Relatedness is defined as $R_{ss} = \frac{N-M}{N} + 0.5 \frac{M}{N}$ where N is number of same-sex twins and M is number of opposite-sex twins. Subsequent columns present variances due to additive genetic effects (A ; column 4); socioeconomic status measure which in our case is years of maternal education at birth (M ; column 5); variance due to shared environment (C ; column 6) and non-shared environment (E ; column 7). Columns (8) to (10) present a moderator term of interest for additive genetic effects (A') with 95% confidence interval. Columns (11) to (13) present a moderator term for shared environment (C') with 95% confidence interval. Columns (14) to (16) present moderator term for non-shared environment (E') with 95% confidence interval. Column (17) presents sample sizes used in estimation. Scarr-Rowe hypothesis requires estimate in column (8) to be positive and statistically significant.

Table S4: Alternative variance components model. Moderator: PCA SES index (excluding zip code income)

Test	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-shared environment)	A' (genetic x SES)	95% CI for A'		C' (common env. x SES)	95% CI for C'		E' (non-shared env. x SES)	95% CI for E'		N
							Lower	Upper		Lower	Upper		Lower	Upper	
Math	3 to 5	0.563	0.150	0.075	0.219	-0.050	-0.071	-0.030	0.029	-0.013	0.070	-0.003	-0.015	0.008	34,270
Reading	3 to 5	0.569	0.151	0.049	0.238	-0.021	-0.041	-0.001	-0.001	-0.046	0.044	-0.002	-0.013	0.009	
Math	6 to 8	0.570	0.175	0.054	0.196	-0.030	-0.063	0.003	-0.008	-0.085	0.068	-0.039	-0.057	-0.022	21,526
Reading	6 to 8	0.540	0.164	0.076	0.214	-0.030	-0.059	-0.001	0.022	-0.035	0.079	0.009	-0.006	0.025	

Notes: See Table S2 for detailed notes. Socioeconomic status measure (*M*) used in this table is based on PCA of: maternal years of education, medicaid paid birth, parents married, father present or father absent.

Table S5: Robustness of alternative variance components model. Moderator: PCA SES index (excluding zip code income)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Test	R _{SS}	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-sh. environment)	A' (genetic x SES)	95% CI for A'		C' (common env. X SES)	95% CI for C'		E' (non-sh. env. X SES)	95% CI for E'		N
Math	0.700	3 to 5	0.717	0.150	0.001	0.140	-0.036	-0.051	-0.021	0.006	-0.097	0.109	-0.006	-0.021	0.009	34,270
Reading		3 to 5	0.686	0.151	0.000	0.171	-0.021	-0.031	-0.012	0.000	-0.052	0.051	0.001	-0.010	0.012	
Math		6 to 8	0.702	0.175	0.000	0.118	-0.033	-0.045	-0.021	0.000	-0.105	0.105	-0.042	-0.057	-0.026	
Reading		6 to 8	0.702	0.164	0.000	0.129	-0.019	-0.031	-0.007	0.000	-0.094	0.095	0.010	-0.006	0.025	
Math	0.725	3 to 5	0.645	0.150	0.035	0.177	-0.044	-0.062	-0.026	0.026	-0.022	0.073	-0.003	-0.017	0.011	34,270
Reading		3 to 5	0.658	0.151	0.005	0.194	-0.021	-0.033	-0.010	0.000	-0.052	0.053	0.000	-0.010	0.010	
Math		6 to 8	0.679	0.175	0.000	0.140	-0.033	-0.044	-0.021	0.000	-0.126	0.127	-0.054	-0.041	-0.027	
Reading		6 to 8	0.645	0.164	0.024	0.162	-0.024	-0.047	-0.001	0.020	-0.051	0.092	0.011	-0.007	0.029	
Math	0.750	3 to 5	0.584	0.150	0.064	0.209	-0.049	-0.069	-0.029	0.029	-0.014	0.071	-0.003	-0.015	0.009	34,270
Reading		3 to 5	0.591	0.151	0.038	0.228	-0.021	-0.040	-0.003	-0.001	-0.048	0.046	-0.002	-0.013	0.010	
Math		6 to 8	0.615	0.175	0.031	0.173	-0.031	-0.062	-0.001	-0.006	-0.098	0.086	-0.040	-0.058	-0.021	
Reading		6 to 8	0.582	0.164	0.055	0.193	-0.028	-0.056	-0.001	0.023	-0.039	0.085	0.010	-0.007	0.027	
Math	0.775	3 to 5	0.533	0.150	0.089	0.234	-0.052	-0.073	-0.031	0.028	-0.011	0.067	-0.004	-0.015	0.007	34,270
Read		3 to 5	0.537	0.151	0.065	0.254	-0.021	-0.042	0.000	-0.002	-0.044	0.040	-0.003	-0.014	0.008	
Math		6 to 8	0.560	0.175	0.060	0.201	-0.030	-0.063	0.004	-0.009	-0.083	0.065	-0.039	-0.056	-0.022	
Reading		6 to 8	0.530	0.164	0.080	0.219	-0.030	-0.059	-0.001	0.022	-0.034	0.077	0.009	-0.006	0.024	
Math	0.800	3 to 5	0.489	0.150	0.110	0.256	-0.053	-0.075	-0.031	0.027	-0.010	0.063	-0.006	-0.016	0.004	34,270
Reading		3 to 5	0.493	0.151	0.087	0.277	-0.020	-0.043	0.002	-0.003	-0.042	0.035	-0.004	-0.014	0.006	
Math		6 to 8	0.513	0.175	0.082	0.225	-0.028	-0.063	0.006	-0.011	-0.075	0.053	-0.039	-0.054	-0.024	
Reading		6 to 8	0.486	0.164	0.102	0.241	-0.031	-0.061	-0.001	0.020	-0.031	0.070	0.007	-0.006	0.021	

Notes: See Table S3 for detailed notes. Socioeconomic status measure (M) used in this table is based on PCA of: maternal years of education, medicaid paid birth, parents married, father present or father absent.

Table S6: Alternative variance components model. Moderator: PCA SES index (including zip code income)

Test	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-shared environment)	A' (genetic x SES)	95% CI for A'		C' (common env. x SES)	95% CI for C'		E' (non-shared env. x SES)	95% CI for E'		N
							Lower	Upper		Lower	Upper		Lower	Upper	
Math	3 to 5	0.573	0.166	0.049	0.219	-0.051	-0.070	-0.032	0.024	-0.020	0.068	-0.004	-0.015	0.008	32,991
Reading	3 to 5	0.567	0.168	0.032	0.240	-0.028	-0.045	-0.012	0.000	-0.042	0.043	0.000	-0.010	0.010	
Math	6 to 8	0.573	0.194	0.029	0.196	-0.034	-0.065	-0.003	-0.007	-0.102	0.089	-0.040	-0.057	-0.023	20,639
Reading	6 to 8	0.521	0.184	0.068	0.220	-0.028	-0.054	-0.002	0.002	-0.048	0.052	0.007	-0.007	0.022	

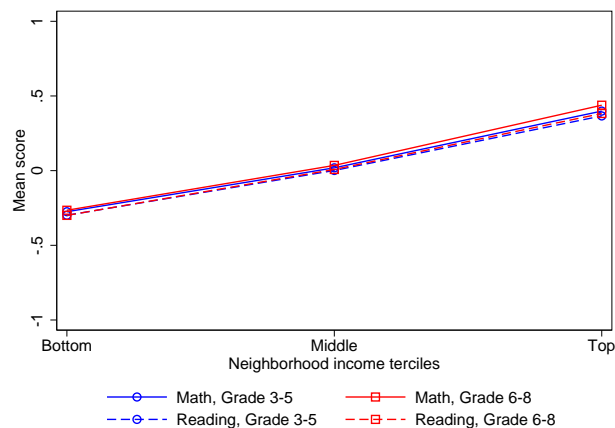
Notes: See Table S2 for detailed notes. Socioeconomic status measure (*M*) used in this table is based on PCA of: maternal years of education, medicaid paid birth, parents married, father present, father absent and median income at zip code of residence at the time of child's birth.

Table S7: Robustness of alternative variance components model. Moderator: PCA SES index (including zip code income)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Test	R _{SS}	Grades	A (additive genetic)	M (SES measure)	C (common environment)	E (non-sh. environment)	A' (genetic x SES)	95% CI for A' Lower	Upper	C' (common env. X SES)	95% CI for C' Lower	Upper	E' (non-sh. env. X SES)	95% CI for E' Lower	Upper	N
Math	0.700	3 to 5	0.692	0.166	0.000	0.151	-0.041	-0.050	-0.031	0.000	-0.060	0.060	-0.004	-0.016	0.008	32,991
Reading		3 to 5	0.659	0.168	0.000	0.181	-0.028	-0.038	-0.019	0.000	-0.045	0.045	0.004	-0.007	0.015	
Math		6 to 8	0.667	0.194	0.000	0.131	-0.036	-0.048	-0.023	0.000	-0.097	0.097	-0.042	-0.057	-0.027	
Reading		6 to 8	0.669	0.184	0.000	0.141	-0.026	-0.038	-0.014	0.000	-0.063	0.063	0.014	-0.001	0.029	
Math	0.725	3 to 5	0.656	0.166	0.010	0.177	-0.043	-0.060	-0.027	0.015	-0.040	0.071	-0.004	-0.017	0.009	32,991
Reading		3 to 5	0.638	0.168	0.000	0.202	-0.028	-0.037	-0.018	0.000	-0.047	0.047	0.002	-0.008	0.012	
Math		6 to 8	0.646	0.194	0.000	0.151	-0.035	-0.047	-0.023	0.000	-0.110	0.110	-0.041	-0.055	-0.028	
Reading		6 to 8	0.624	0.184	0.016	0.168	-0.027	-0.045	-0.009	0.003	-0.058	0.064	0.011	-0.004	0.027	
Math	0.750	3 to 5	0.594	0.166	0.039	0.208	-0.049	-0.068	-0.031	0.023	-0.022	0.069	-0.004	-0.016	0.008	32,991
Reading		3 to 5	0.590	0.168	0.022	0.229	-0.028	-0.043	-0.013	0.000	-0.044	0.044	0.001	-0.010	0.011	
Math		6 to 8	0.618	0.194	0.006	0.173	-0.035	-0.055	-0.015	-0.004	-0.124	0.116	-0.041	-0.056	-0.026	
Reading		6 to 8	0.563	0.184	0.047	0.200	-0.028	-0.052	-0.004	0.003	-0.051	0.057	0.009	-0.006	0.025	
Math	0.775	3 to 5	0.543	0.166	0.064	0.234	-0.052	-0.072	-0.033	0.024	-0.017	0.065	-0.005	-0.016	0.006	32,991
Read		3 to 5	0.536	0.168	0.048	0.256	-0.029	-0.047	-0.010	0.000	-0.040	0.040	-0.001	-0.011	0.010	
Math		6 to 8	0.563	0.194	0.034	0.202	-0.034	-0.066	-0.001	-0.007	-0.098	0.084	-0.040	-0.057	-0.023	
Reading		6 to 8	0.511	0.184	0.073	0.226	-0.028	-0.055	-0.001	0.002	-0.047	0.051	0.007	-0.007	0.021	
Math	0.800	3 to 5	0.500	0.166	0.085	0.257	-0.054	-0.075	-0.033	0.023	-0.015	0.061	-0.006	-0.016	0.004	32,991
Reading		3 to 5	0.491	0.168	0.070	0.278	-0.028	-0.048	-0.008	-0.001	-0.038	0.036	-0.002	-0.011	0.008	
Math		6 to 8	0.516	0.194	0.057	0.225	-0.033	-0.067	0.001	-0.008	-0.084	0.067	-0.040	-0.055	-0.025	
Reading		6 to 8	0.469	0.184	0.094	0.247	-0.028	-0.056	0.000	0.001	-0.045	0.046	0.005	-0.007	0.018	

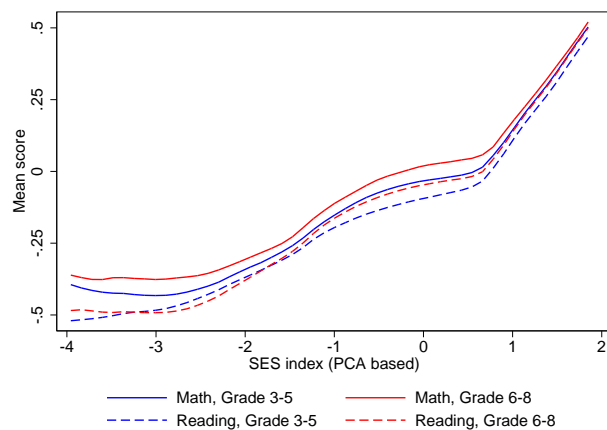
Notes: See Table S3 for detailed notes. Socioeconomic status measure (M) used in this table is based on PCA of: maternal years of education, medicaid paid birth, parents married, father present, father absent and median income at zip code of residence at the time of child's birth.

Figure S1: Neighborhood income at birth and average achievement test score for combined twin and sibling pairs sample



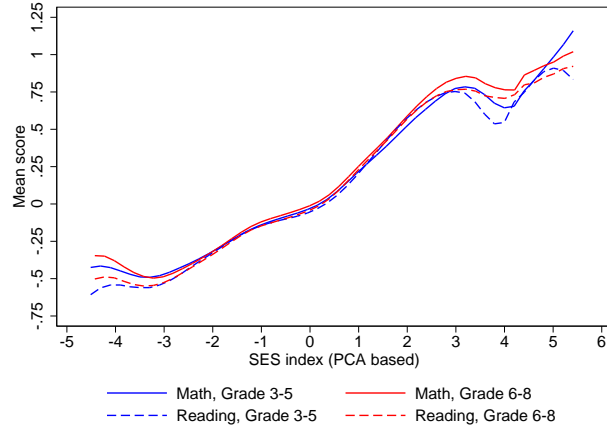
Notes: This figure plots means of gender-standardized test scores in mathematics and reading over zip-code level neighborhood income terciles at birth (separated by age group and test type). Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002.

Figure S2: SES index (PCA excluding zip code income) at birth and average achievement test score for combined twin and sibling pairs sample



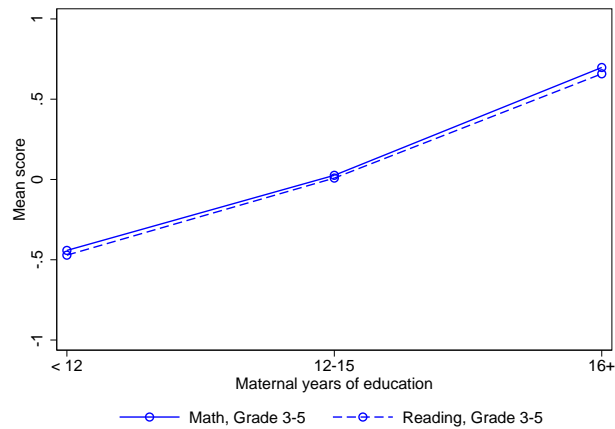
Notes: This figure plots means (smoothed polynomial) of gender-standardized test scores in mathematics and reading over SES index based on maternal years of education, medicaid paid birth, mother married, father present and father absent. Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002.

Figure S3: SES index (PCA including zip code income) at birth and average achievement test score for combined twin and sibling pairs sample



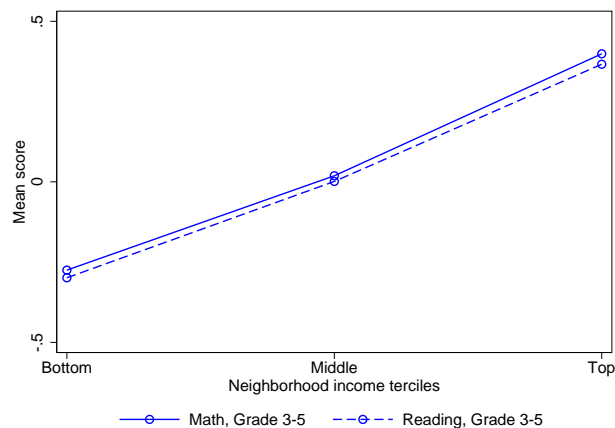
Notes: This figure plots means (smoothed polynomial) of gender-standardized test scores in mathematics and reading over SES index based on maternal years of education, medicaid paid birth, mother married, father present, father absent and median zip-code level neighborhood income at birth. Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002.

Figure S4: Maternal years of education and average Stanford achievement test score for combined twin and sibling pairs sample



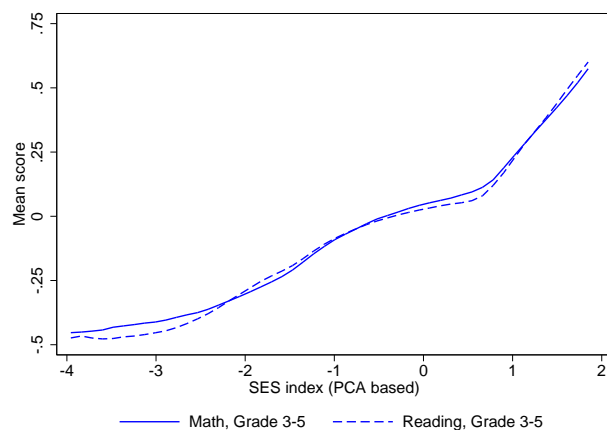
Notes: This figure plots means of gender-standardized Stanford achievement test scores in mathematics and reading over maternal years of education (separated by test type). Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. Stanford achievement test is only available for grades 3 to 5.

Figure S5: Neighborhood income at birth and average Stanford achievement test score for combined twin and sibling pairs sample



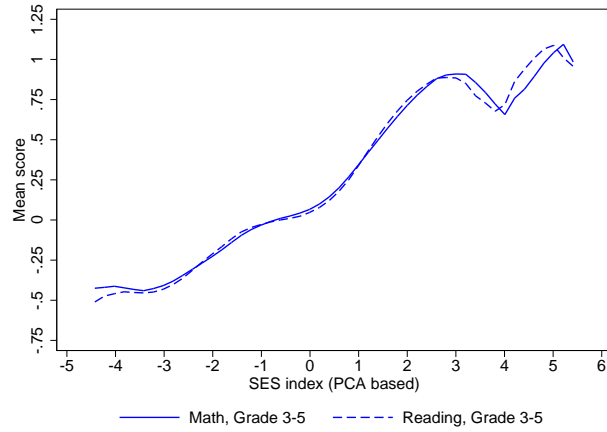
Notes: This figure plots means of gender-standardized Stanford achievement test scores in mathematics and reading over zip-code level neighborhood income terciles at birth (separated by test type). Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. Stanford achievement test is only available for grades 3 to 5.

Figure S6: SES index (PCA excluding zip code income) at birth and average Stanford achievement test score for combined twin and sibling pairs sample



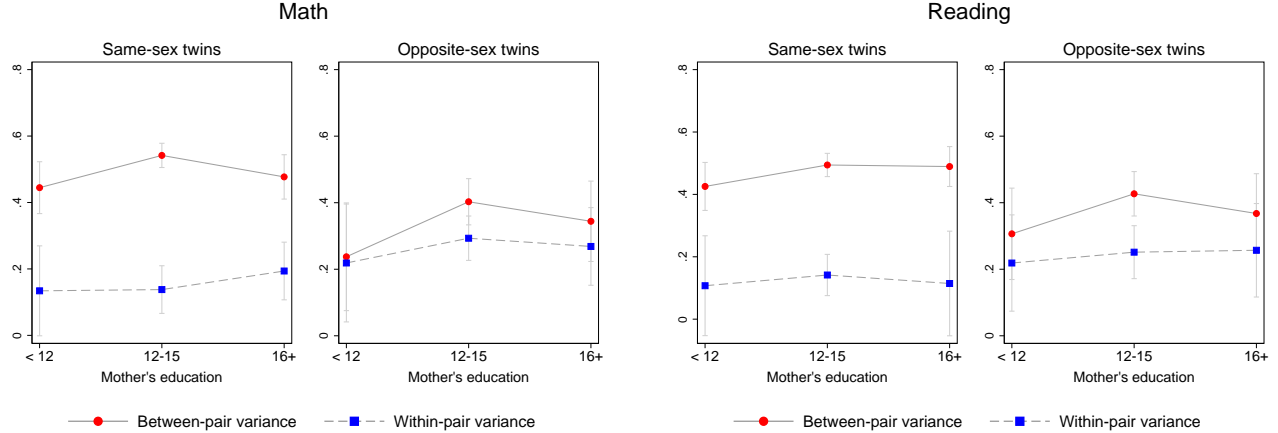
Notes: This figure plots means (smoothed polynomial) of gender-standardized Stanford achievement test scores in mathematics and reading over SES index based on maternal years of education, medicaid paid birth, mother married, father present and father absent (separated by test type). Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. Stanford achievement test is only available for grades 3 to 5.

Figure S7: SES index (PCA including zip code income) at birth and average Stanford achievement test score for combined twin and sibling pairs sample



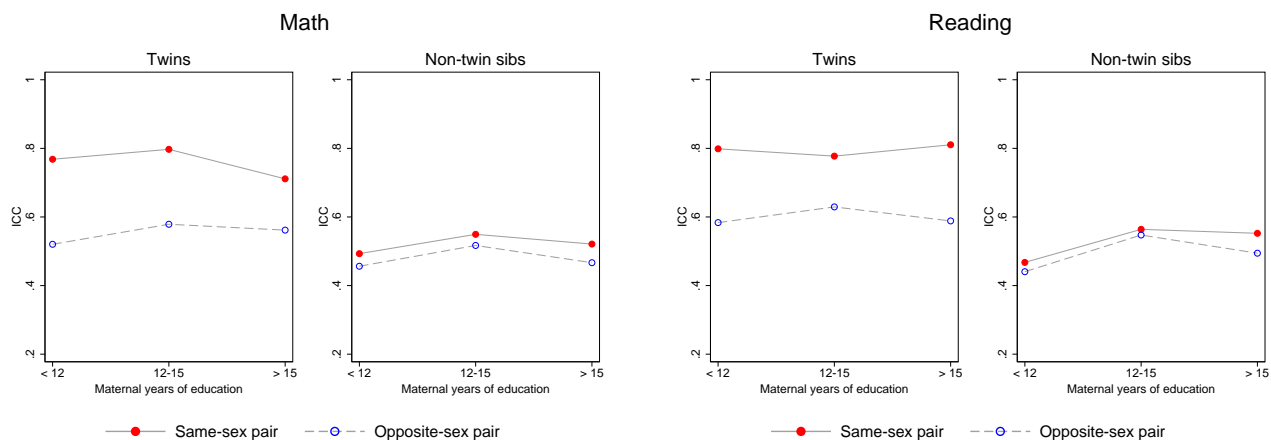
Notes: This figure plots means (smoothed polynomial) of gender-standardized Stanford achievement test scores in mathematics and reading over SES index based on maternal years of education, medicaid paid birth, mother married, father present, father absent and median zip-code level neighborhood income at birth (separated by test type). Sample includes all twin pairs and closely spaced sibling pairs with available test scores. Closely spaced sibling pair is defined as two siblings having the same mother for whom the distance in months between births is the smallest among births to this mother between 1994 and 2002. Stanford achievement test is only available for grades 3 to 5.

Figure S8: Between- and within-pair variance in Stanford achievement test for same-sex and opposite sex twins



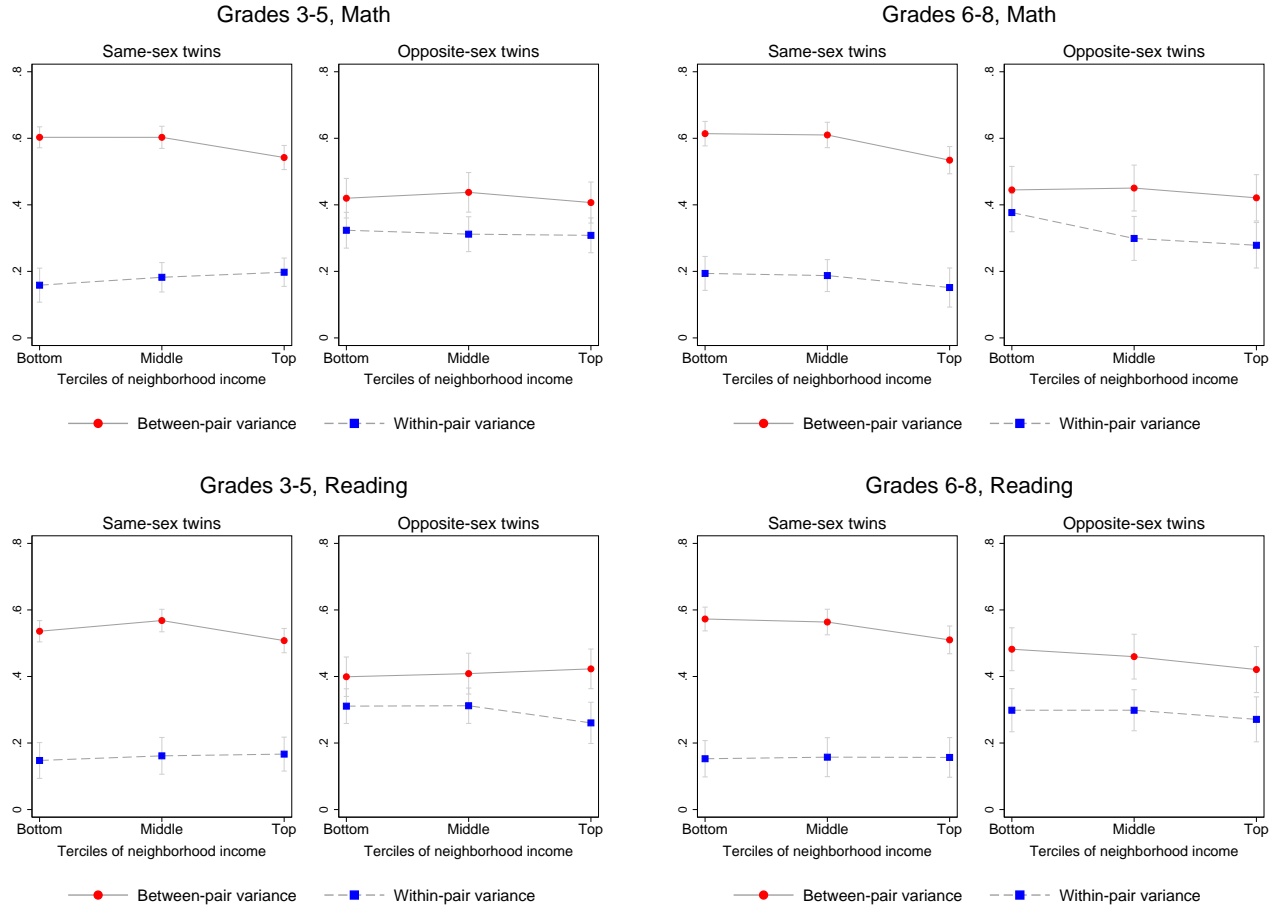
Notes: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 Stanford mathematics achievement test. The right-hand side figures present identical results for reading test. For estimation details see Figure 3.

Figure S9: Intraclass correlations for same-sex and opposite-sex twin and non-twin sibling pairs on Stanford achievement test



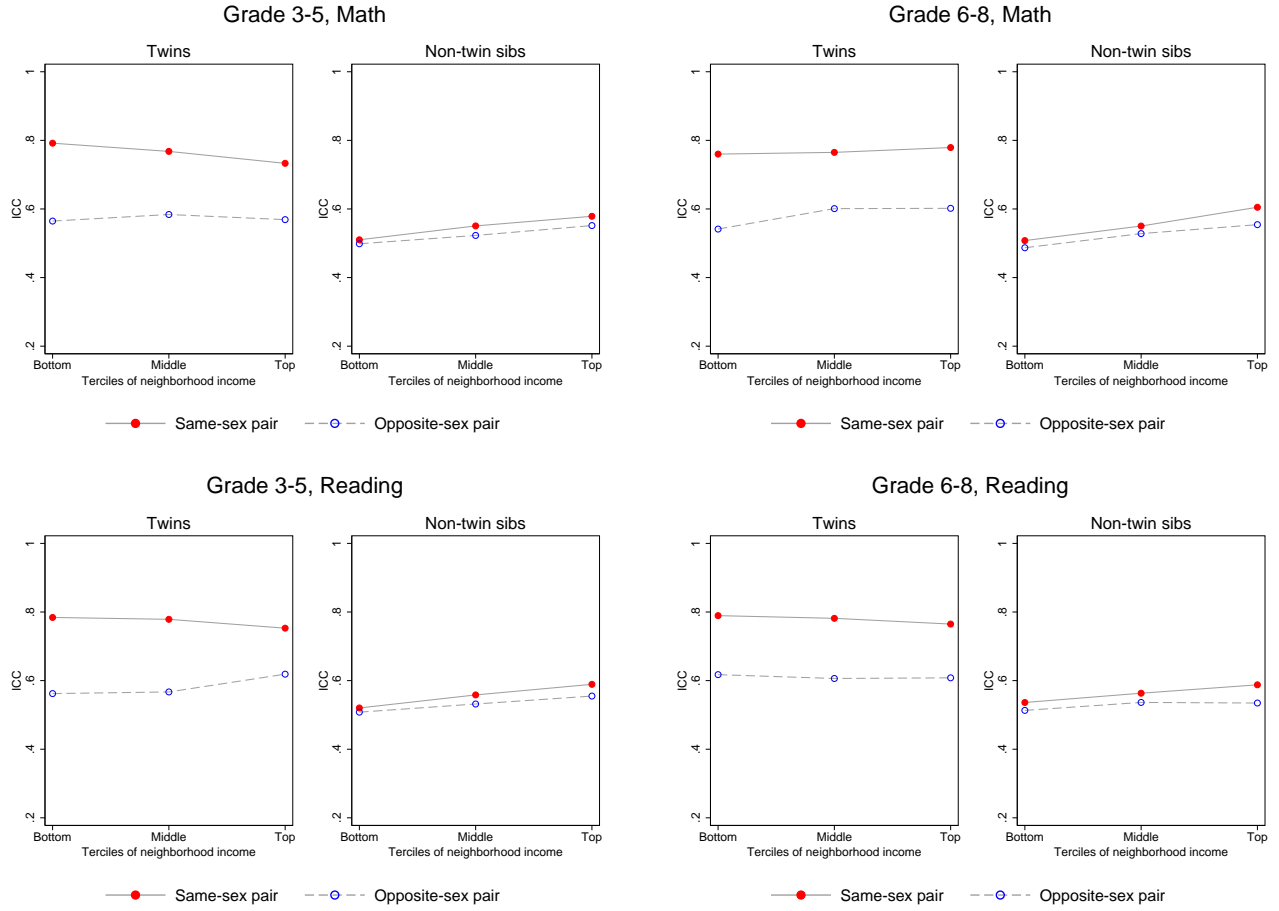
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 Stanford achievement test. The right-hand side figures present identical results for reading test. For estimation details see Figure 4.

Figure S10: Between- and within-pair variance in achievement test scores for same-sex and opposite-sex twins. Neighborhood income SES measure.



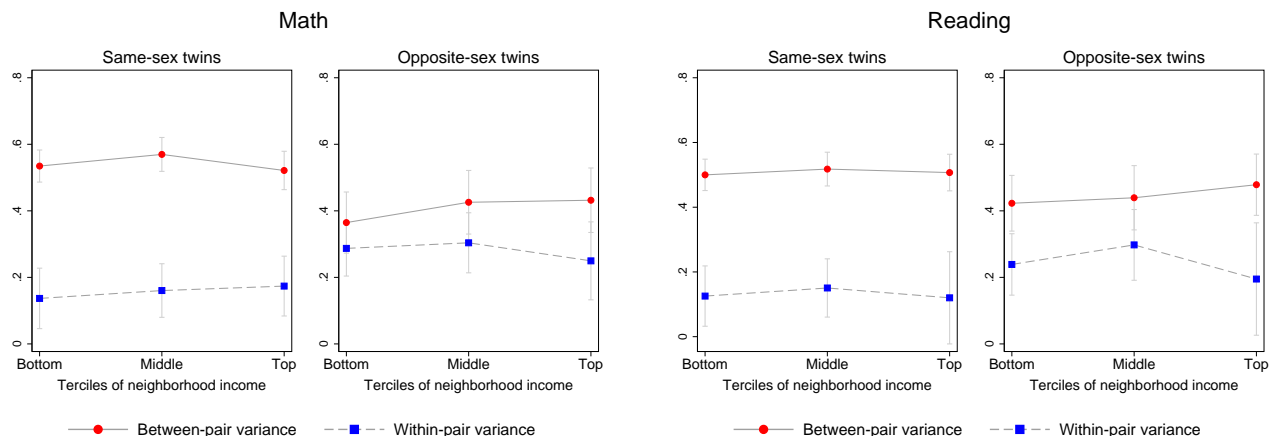
Note: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on zip-code level neighborhood income. For estimation details see Figure 3.

Figure S11: Intraclass correlation for same-sex and opposite-sex twin and non-twin sibling pairs.
Neighborhood income SES measure.



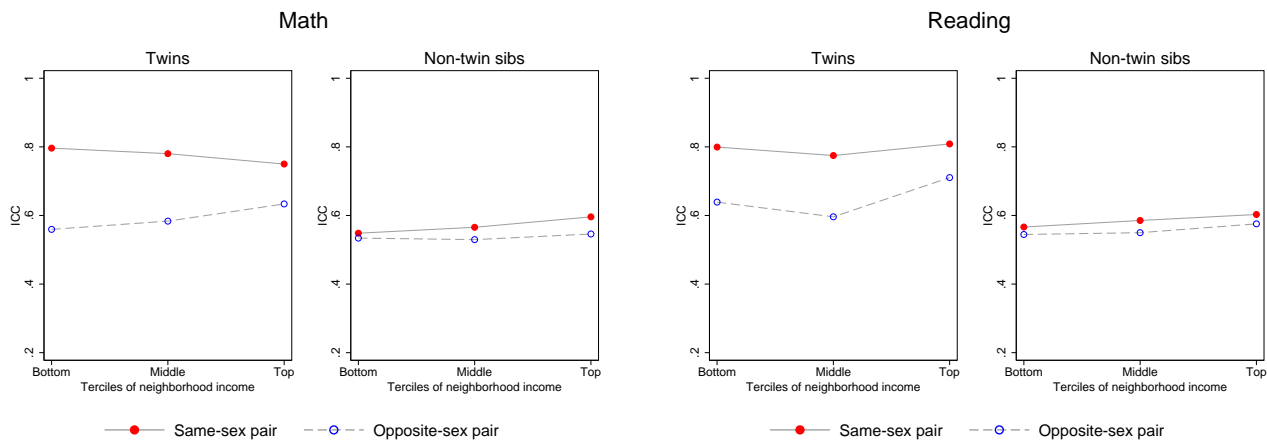
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on zip-code level neighborhood income. For estimation details see Figure 4.

Figure S12: Between- and within-pair variance in Stanford achievement test for same-sex and opposite-sex twins. Neighborhood income SES measure.



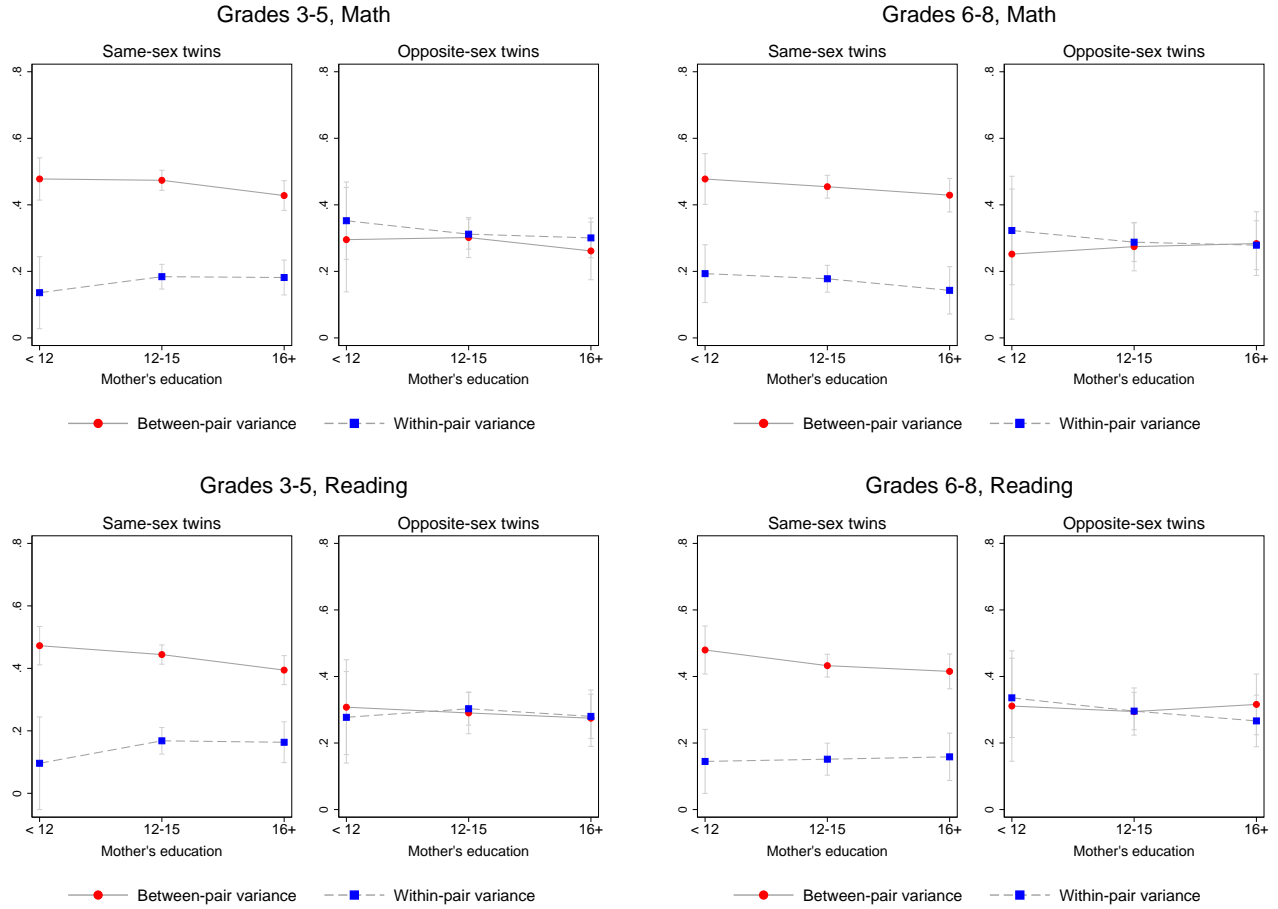
Notes: The left-hand side figures presents between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 Stanford mathematics achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on zip-code level neighborhood income. For estimation details see Figure 3.

Figure S13: Intraclass correlations for same-sex and opposite-sex twin and non-twin sibling pairs on Stanford achievement test. Neighborhood income SES measure.



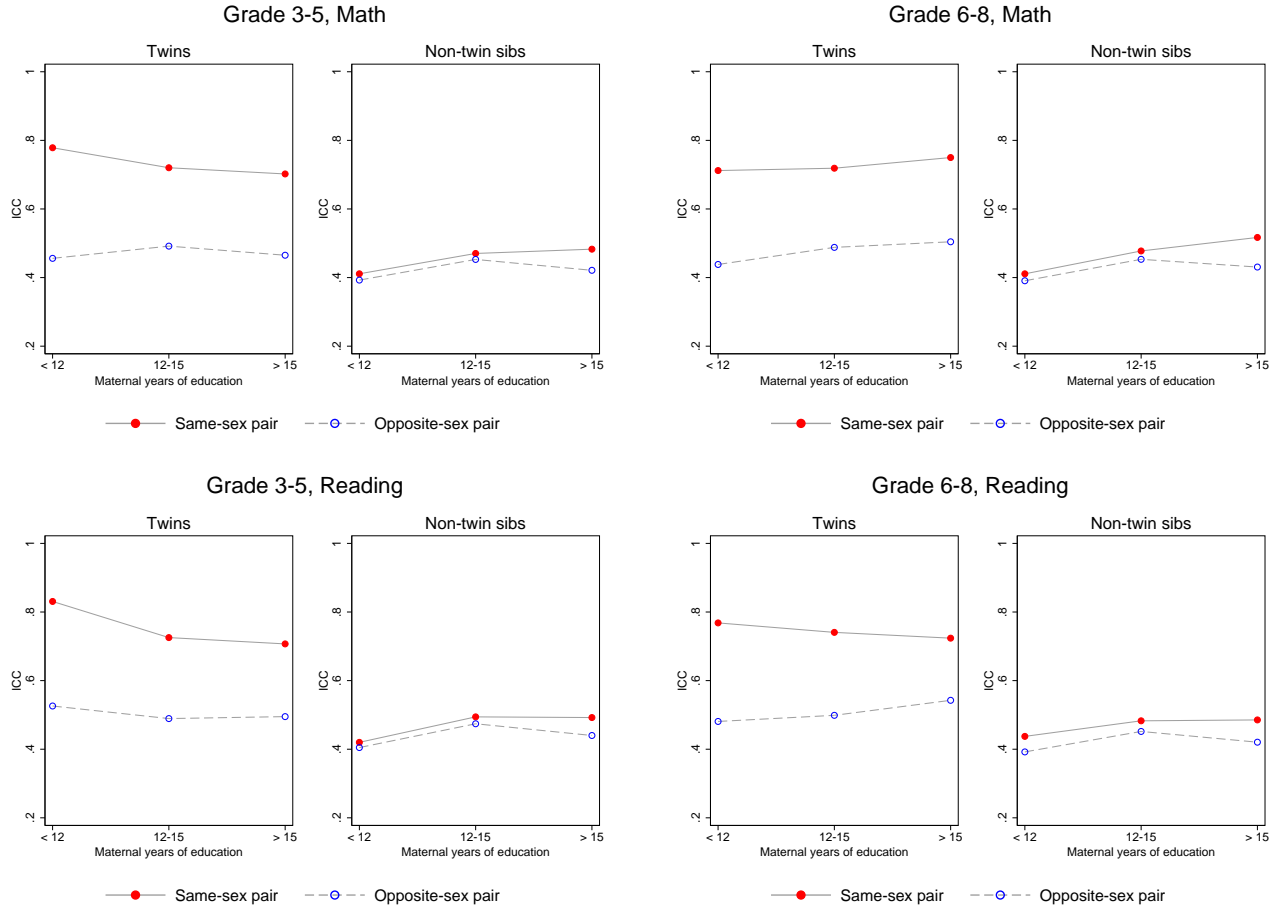
Note: The left-hand side figures presents intraclass correlations among same-sex and opposite-sex pairs of twins and non-twin siblings in grades 3 to 5 Stanford mathematics achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on zip-code level neighborhood income. For estimation details see Figure 4.

Figure S14: Between- and within-pair variance in achievement test scores for same-sex and opposite-sex twins. White mothers



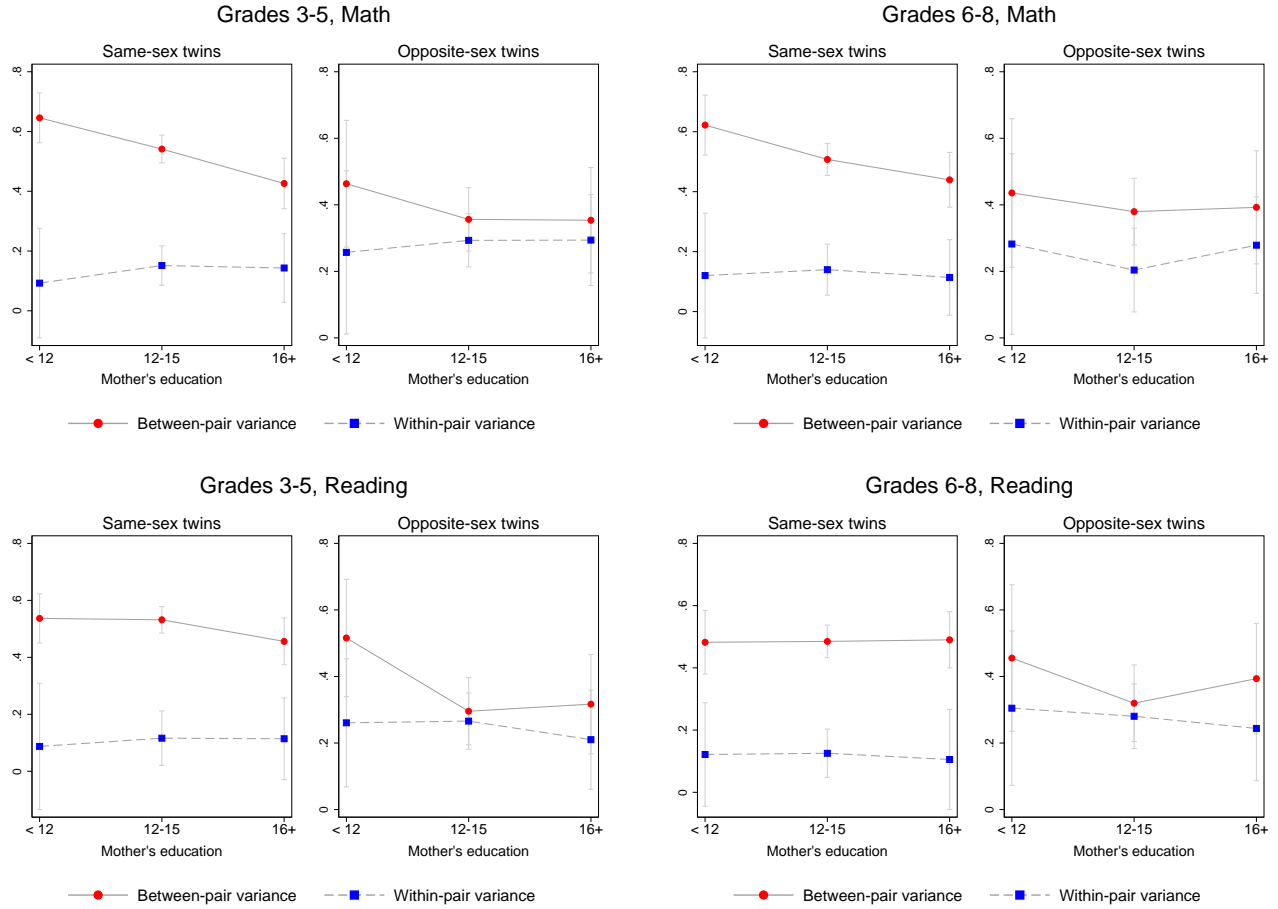
Note: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on maternal years of education. Sample of white mothers. For estimation details see Figure 3.

Figure S15: Intraclass correlation for same-sex and opposite-sex twin and non-twin sibling pairs.
White mothers.



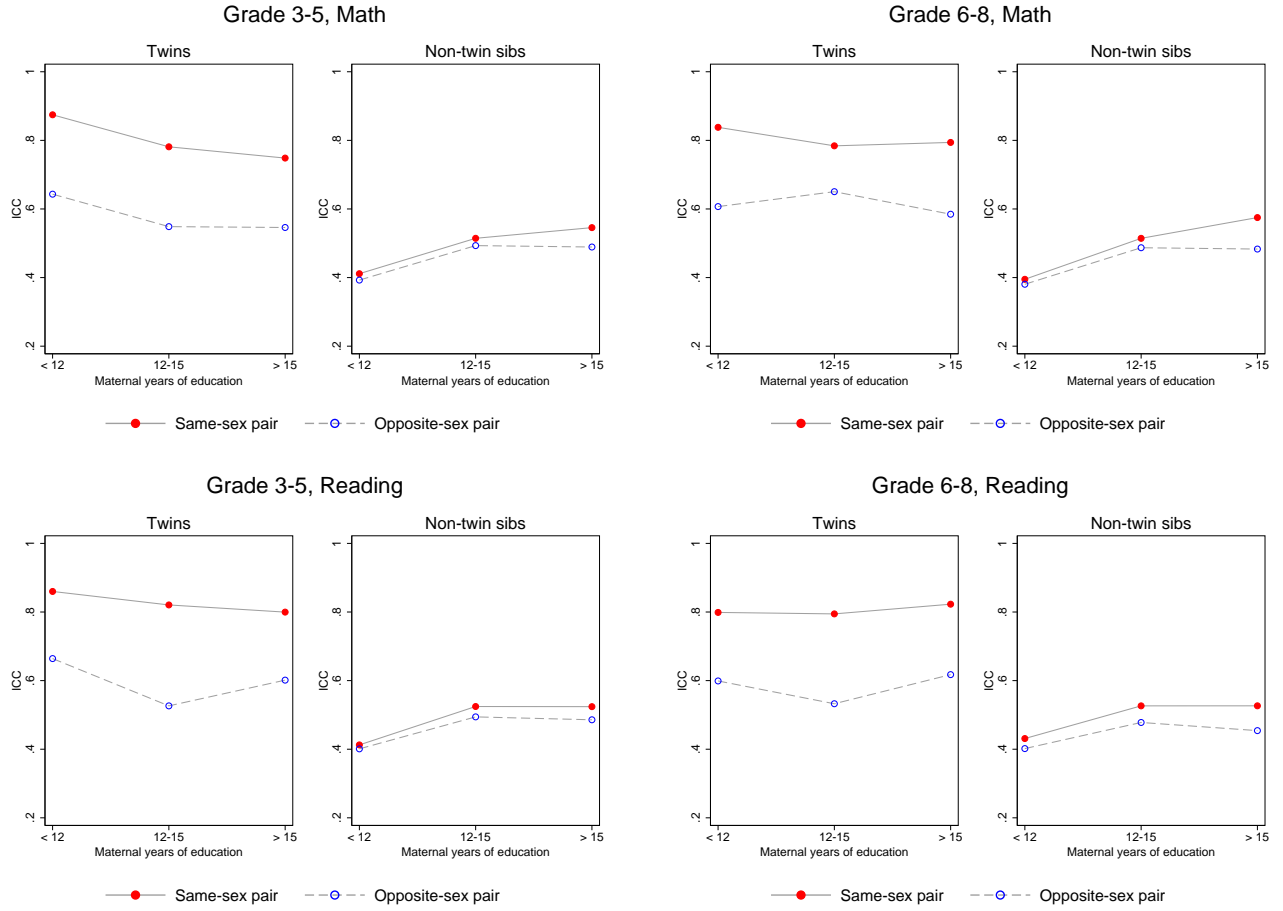
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on maternal years of education. Sample of white mothers. For estimation details see Figure 4.

Figure S16: Between- and within-pair variance in achievement test scores for same-sex and opposite-sex twins. Mothers up to age 30.



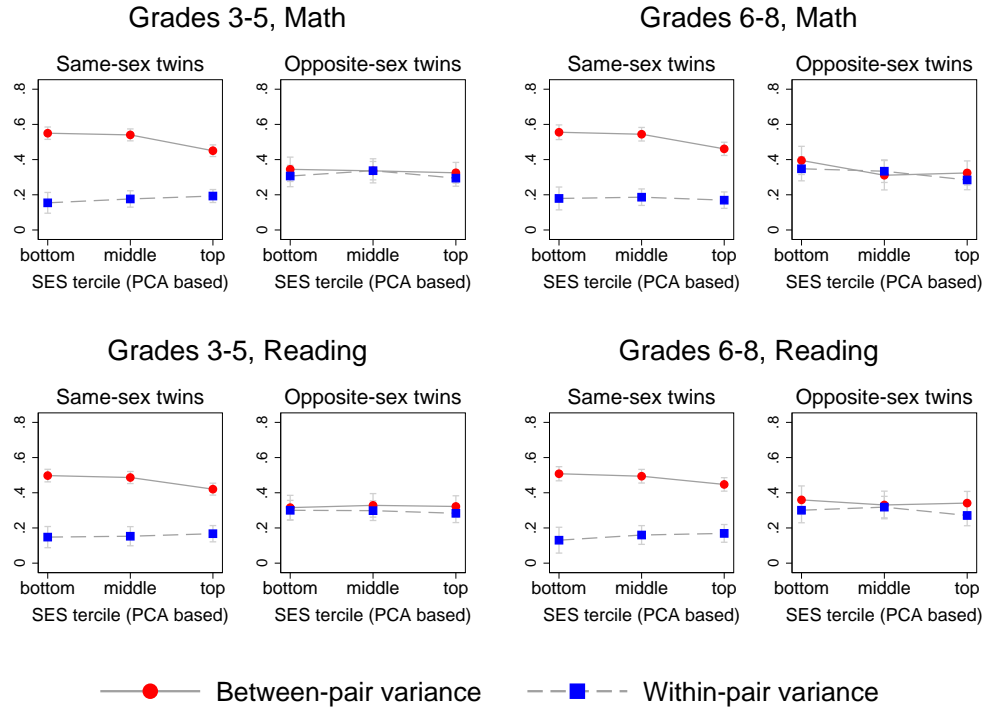
Note: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on maternal years of education. Sample of mothers up to the age of 30. For estimation details see Figure 3.

Figure S17: Intraclass correlation for same-sex and opposite-sex twin and non-twin sibling pairs.
Mothers up to age 30.



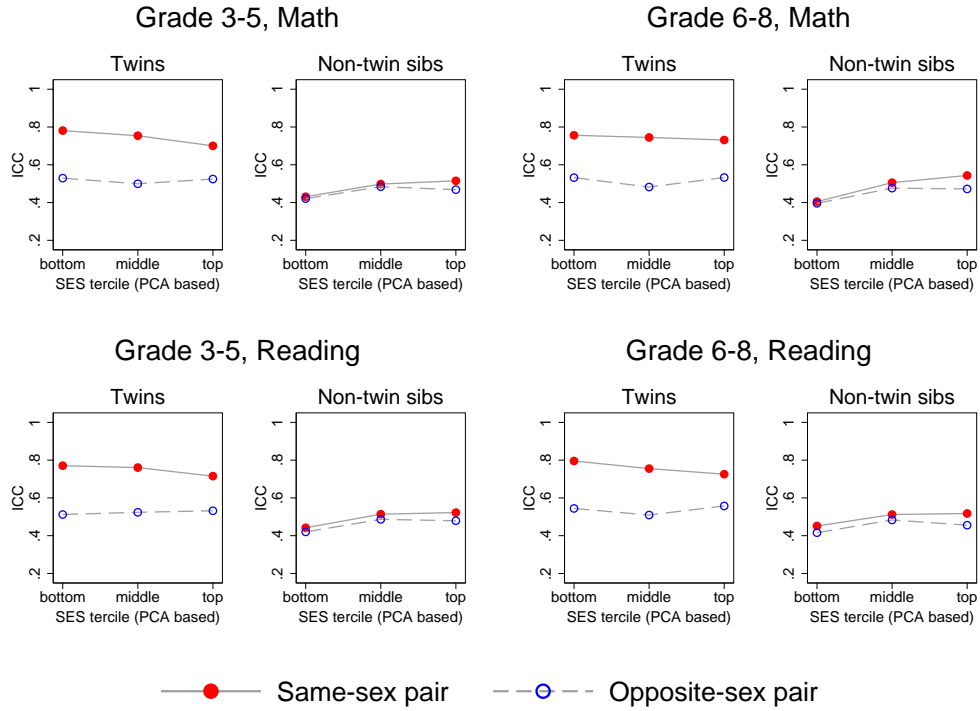
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on maternal years of education. Sample of mothers up to age of 30. For estimation details see Figure 4.

Figure S18: Between- and within-pair variance in achievement test scores for same-sex and opposite-sex twins. SES terciles based on PCA (excluding zip code income).



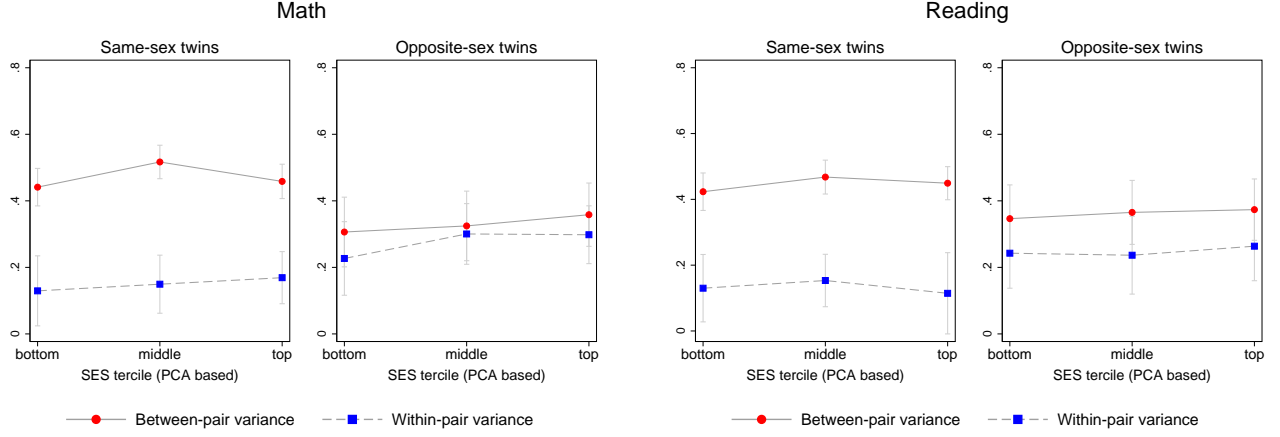
Note: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present and father absent. For estimation details see Figure 3.

Figure S19: Intraclass correlation for same-sex and opposite-sex twin and non-twin sibling pairs.
SES terciles based on PCA (excluding zip code income).



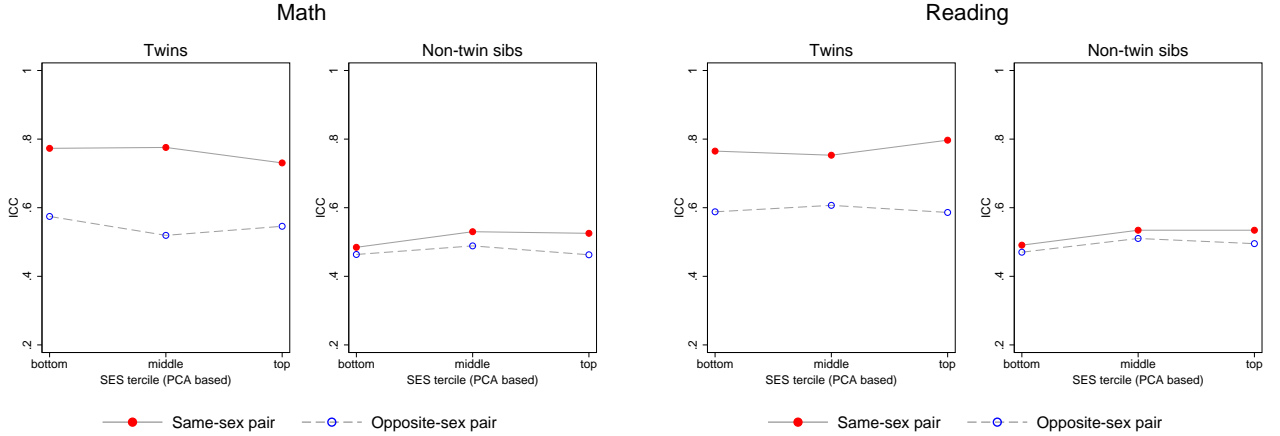
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present and father absent. For estimation details see Figure 4.

Figure S20: Between- and within-pair variance in Stanford achievement test for same-sex and opposite sex twins. SES terciles based on PCA (excluding zip code income).



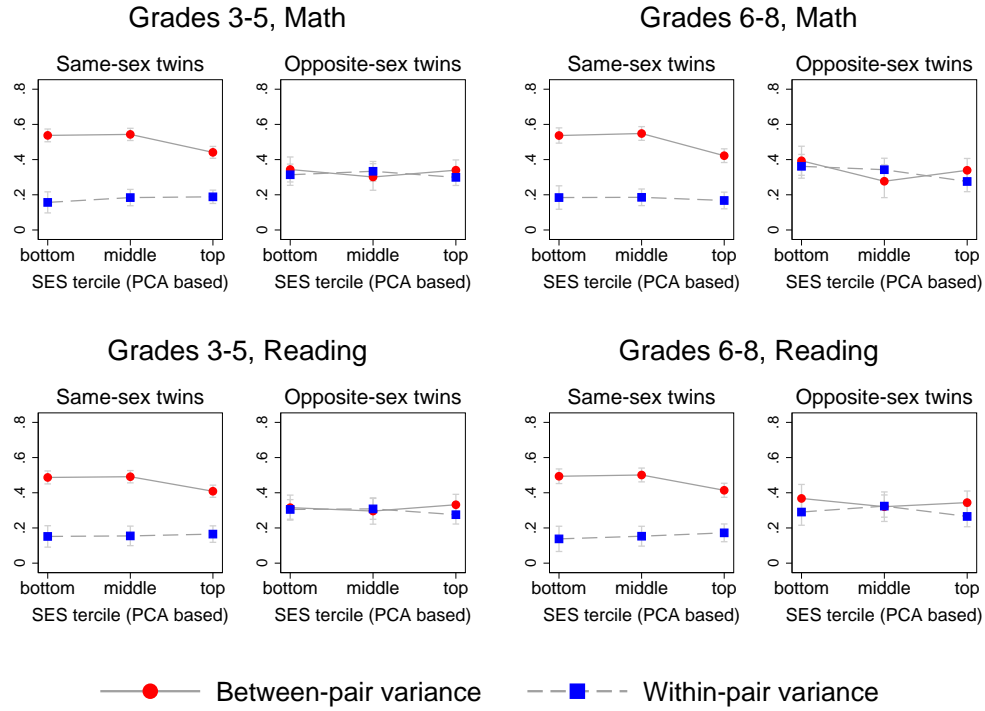
Notes: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 Stanford mathematics achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present and father absent. For estimation details see Figure 3.

Figure S21: Intraclass correlations for same-sex and opposite-sex twin and non-twin sibling pairs on Stanford achievement test. SES terciles based on PCA (excluding zip code income).



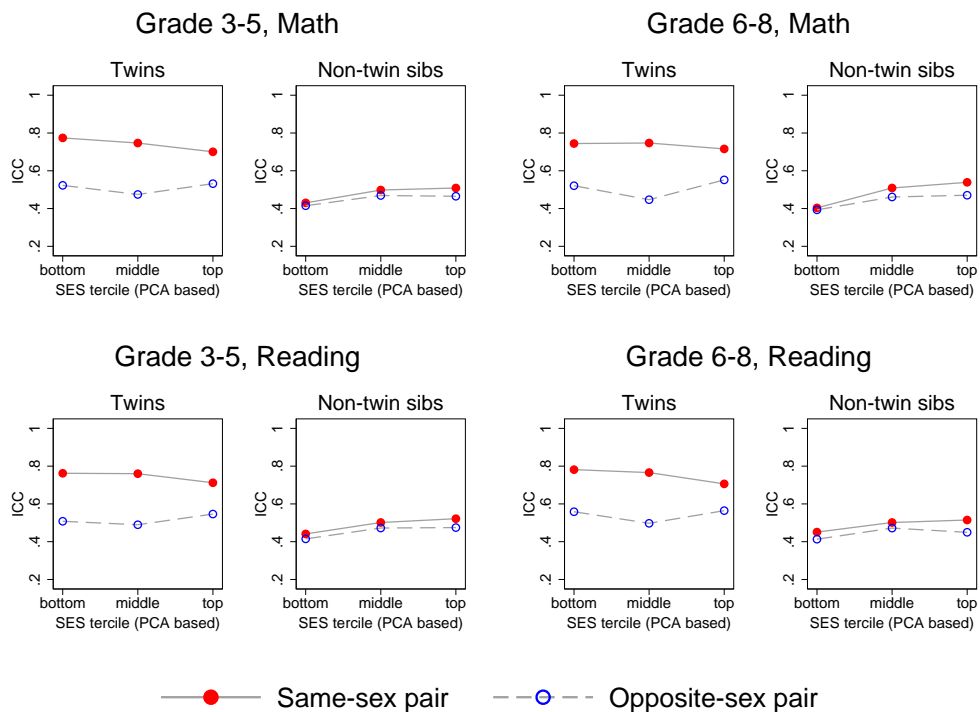
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 Stanford achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present and father absent. For estimation details see Figure 4.

Figure S22: Between- and within-pair variance in achievement test scores for same-sex and opposite-sex twins. SES terciles based on PCA (including zip code income).



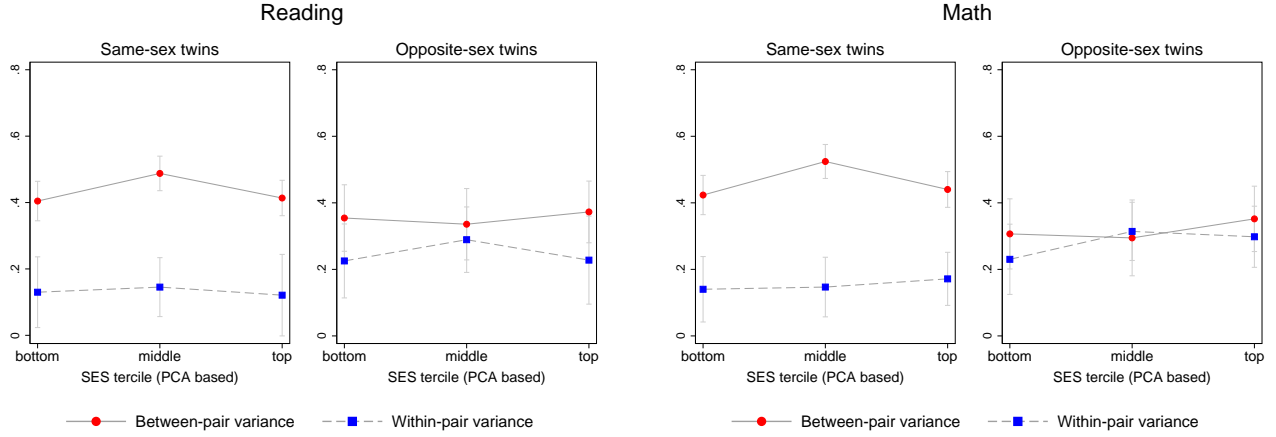
Note: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present, father absent and median zip code income of residence at the time of child birth. For estimation details see Figure 3.

Figure S23: Intraclass correlation for same-sex and opposite-sex twin and non-twin sibling pairs.
SES terciles based on PCA (including zip code income).



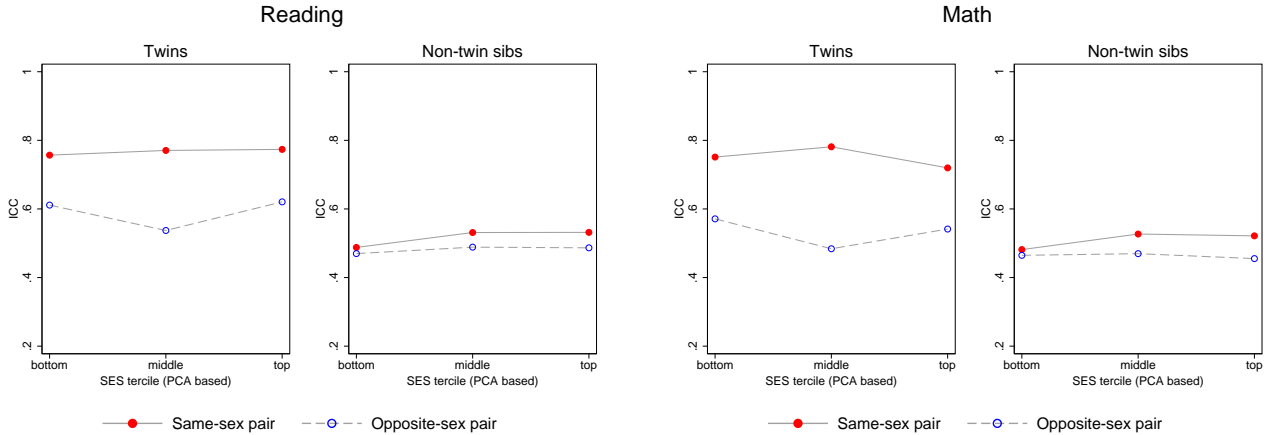
Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 mathematics (top) and reading (bottom) FCAT test scores. The right-hand side figures present identical results for grades 6 to 8. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present, father absent and median zip code income of residence at the time of child birth. For estimation details see Figure 4.

Figure S24: Between- and within-pair variance in Stanford achievement test for same-sex and opposite sex twins. SES terciles based on PCA (including zip code income).



Notes: The left-hand side figures present between- and within-pair variance among same-sex and opposite-sex twins in grades 3 to 5 Stanford mathematics achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present, father absent and median zip code income of residence at the time of child birth. For estimation details see Figure 3.

Figure S25: Intraclass correlations for same-sex and opposite-sex twin and non-twin sibling pairs on Stanford achievement test. SES terciles based on PCA (including zip code income).



Note: The left-hand side figures present intraclass correlation coefficients among same-sex and opposite-sex pairs of twins and closely spaced siblings in grades 3 to 5 Stanford achievement test. The right-hand side figures present identical results for reading test. Socioeconomic status is measured based on index constructed using PCA where we include maternal years of education, medicaid paid birth, parents married, father present, father absent and median zip code income of residence at the time of child birth. For estimation details see Figure 4.