LS
7/26
1:15 PM

Does the Sibling Correlation in Economic Status Vary Across Families and Sibling Pairs?

Gerald S. Oettinger*

Department of Economics

University of Texas

April 29, 1999

(Preliminary and Incomplete)

Abstract

This paper uses data from the National Longitudinal Survey of Youth to estimate sibling correlations in educational attainment and cognitive ability and to investigate whether these correlations vary systematically with family or sibling pair characteristics. Similar to previous authors, I estimate raw sibling correlations of .46-.57 for educational attainment and .62-.67 for cognitive ability and I find that observable differences in family background between families can account for less than half of these correlations. An analysis of the residual sibling correlations in status shows that the degree of sibling resemblance varies systematically with characteristics of the family and the sibling pair. In particular, sibling correlations in status appears to be U-shaped in father's education (a proxy for family permanent income), implying that siblings from middle income families have less similar outcomes than siblings from both rich and poor families. There is also some evidence that the sibling correlation in status is smaller for siblings far apart in age, which suggests that time-varying components of family environment have a non-trivial impact on eventual socioeconomic status.

JEL Classification: I21, J24, J62

*Address: Department of Economics, University of Texas, Austin, TX 78712; Phone: (512)-475-8523; Fax: (512)-471-3510. Email: oettinge@eco.utexas.edu. I thank Alex Minicozzi for helpful discussions and seminar participants at the University of Texas for comments. All errors are mine.

1 Introduction

A long line of research documents that the correlation between siblings in adult economic status is positive and large. For example, typical estimates of the sibling correlation in educational attainment are around 0.5 (Behrman, Taubman, and Wales 1977; Chamberlain and Griliches 1977; Olneck 1977; Behrman and Taubman 1989; Solon, Page, and Duncan 1997) and typical estimates for the correlation in permanent (log) earnings are around 0.4 (Altonji and Dunn 1991; Solon, et. al. 1991).[1] With nearly half of the overall variation in economic status accounted for by factors that siblings share in common, it is hard not to conclude that the sum total of shared family background — including common genetic influences, family permanent income, and parental values and preferences — plays a crucial role in determining eventual economic status. However, this same evidence implies that roughly half of the overall variation in economic status occurs within families, and the source of this within-family variation is much less clear.

Economic models of intrahousehold allocation imply that the family also might be the source of much of the within-family variation in educational attainment and earnings.[2] In particular, differences between siblings in endowed abilities will tend to generate sibling differences in economic outcomes, which parents might either reinforce or compensate through their allocation of human capital investments among their children. Several authors (Behrman, Pollak, and Taubman 1982; Behrman, Pollak, and Taubman 1986; Behrman and Taubman 1986; Behrman, Rosenzweig, and Taubman 1994) have investigated whether parents compensate or reinforce sibling differences in endowments. The evidence generally suggests that parents reinforce endowment differences (i.e., invest more in more able children) but that reinforcement is tempered by concern for equity in child outcomes (i.e.,

[1] Solon (1998) surveys the empirical literature on sibling correlations and intergenerational correlations in earnings.

[2] This class of models views parents as making optimal human capital investments in (and, perhaps, financial transfers to) their children, given parental preferences over child outcomes and the constraints imposed by the children's abilities, the earnings production function, and the resources parents are willing to devote to aggregate human capital investment (and transfers). Behrman (1997) surveys the theoretical and empirical contributions to this literature.

investments are not allocated to maximize the *sum* of children's earnings).

There is no reason to expect, however, that family allocation decisions and consequently the amount of within-family variation in child outcomes will be the same for all types of families and all types of sibling groups. For example, suppose that parents care about the wealth of their children and are equally concerned about all of their children, as in Becker and Tomes (1976) and Becker (1981). Parents who devote sufficiently many resources to their children invest in each child's human capital until the marginal return on human capital equals the return on market investments and then fully compensate any resulting earnings differences with financial transfers. In contrast, poorer parents only invest in their children's human capital (i.e., no financial transfers are made) and therefore the allocation of the human capital investments depends on the parents' aversion to inequality among their children as well as on efficiency considerations. Thus, the Becker-Tomes model suggests that, for any given difference in sibling endowments, the sibling difference in human capital and earnings is likely to increase with family wealth.[3]

In addition, if parents cannot fully smooth fluctuations in household income and other components of family environment, siblings born at different times will be raised in different conditions. This suggests that the extent of sibling resemblance in economic status might depend on the sibling age difference. For example, since siblings further apart in age are likely to experience both larger differences in "average" household environment and fewer common transitory shocks to household environment, the sibling resemblance in economic outcomes might decrease with the sibling age difference.

This discussion suggests that the degree of sibling resemblance in economic status is likely to vary systematically with family and sibling pair characteristics. Moreover, the observed empirical patterns potentially can offer clues about the nature of parental preferences over child outcomes, the technology for producing these outcomes, and the ability of

[3]This conclusion relies on the standard assumption that the marginal return on human capital investment is higher for more able siblings than for less able siblings at all levels of human capital investment.

families to smooth long-term trends and short-term fluctuations in household environment. It is perhaps surprising, then, that there has been almost no investigation of whether the sibling correlation in economic status varies with family and sibling pair characteristics. The main contribution of the present paper is to provide evidence on this question. The empirical approach builds on recent analyses of sibling correlations in status. However, in addition to computing sibling correlation estimates analogous to those in the literature, the present paper also estimates models in which family and sibling pair characteristics are allowed to influence the sibling correlation in status.

The empirical analysis reveals that the sibling correlation in economic status is U-shaped in father's education. More specifically, siblings with fathers who have exactly a high school degree have less similar economic outcomes than siblings with fathers who either went to college or did not finish high school. Since father's education is a proxy for family permanent income, this result suggests that siblings from middle income families achieve less similar outcomes than do siblings from both rich and poor families. While not predicted by the standard stylized models of intrafamily allocation, this finding can be explained by plausible extensions of these models. In addition, the data show that the sibling correlation in educational attainment is lower for siblings who are far apart in age. This finding suggests that families cannot fully smooth out changes in the time-varying components of household environment and that these elements of the household environment have a nonneglible impact on a child's eventual economic status.

The remainder of the paper proceeds as follows. Section 2 develops the empirical methodology used in the analysis. Section 3 briefly describes the data. Section 4 presents the empirical results and discusses what they suggest about how families allocate resources among children. Section 5 concludes.

2 Empirical Methodology

Suppose one has survey data on some time-invariant measure of individual economic status, such as completed educational attainment, for multiple siblings from each of a large number of families. A standard empirical model for economic status is

$$Y_{ij} = X_{ij}\beta + \epsilon_{ij}, \tag{1}$$

where $i = 1, \ldots, N$ indexes families, $j = 1, \ldots J_i$ indexes the observed siblings from family i (with $J_i \geq 2$), Y_{ij} is a measure of economic status, X_{ij} is a vector of exogenous influences on status, and ϵ_{ij} is a random error. Further, let ϵ_{ij} be decomposed into two orthogonal components, $\epsilon_{ij} = \alpha_i + u_{ij}$, with α_i capturing all of the unobserved determinants of status that are shared by siblings in the same family and u_{ij} representing all of the unobserved sibling-specific determinants of status.

Researchers have often been interested in computing the sibling correlation in economic status as a summary measure of the importance of family background in determining economic status, and obtaining such estimates is the starting point of the present analysis as well. Assuming that both α_i and u_{ij} have constant variances, the sibling correlation in economic status (after controlling for X) is $\rho = \sigma_\alpha^2/\sigma_\epsilon^2 = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_u^2)$. Since interest typically centers on measuring the proportion of the overall variation in economic status that is explained by *all* dimensions of family background, the vector X typically includes a very sparse set of variables (e.g., cohort effects).[4]

An estimate of ρ can be constructed from estimates of σ_α^2 and σ_ϵ^2, which in turn can be

[4]Even in this case, ρ does not isolate the contributions of family background since some of the common variation in sibling outcomes is due to shared influences outside of the family (e.g., neighborhood effects) while some of the individual-specific variation originates within the family (e.g., sibling-specific genetic endowments and parental investments).

obtained from sample moments of residuals from (1). In particular, an estimator of σ_ϵ^2 is

$$\hat{\sigma}_\epsilon^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J_i} w_i \hat{\epsilon}_{ij}^2}{\sum_{i=1}^{N}\sum_{j=1}^{J_i} w_i}, \tag{2}$$

where $\hat{\epsilon}_{ij}$ are regression residuals from (1) and w_i is the weight given to family i in computing the sample average of the squared residuals. The weighting factor, w_i, equals 1 if all families receive the same weight and equals $1/p_i$, where p_i is family i's probability of being surveyed, if one accounts for between-family differences in survey probabilities. The estimate $\hat{\sigma}_\epsilon^2$ can be obtained easily as the mean squared error from the (weighted) regression in (1).

Solon, Page, and Duncan (1997) note that a general estimator for σ_α^2, the sibling covariance in economic status, is

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^{N} \tilde{w}_i \{[\sum_{j=1}^{J_i-1}\sum_{j'=j+1}^{J_i} \hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}]/[J_i(J_i-1)/2]\}}{\sum_{i=1}^{N}\tilde{w}_i}, \tag{3}$$

where $\tilde{w}_i$ again represents the weight assigned to family i in the forming the sample average of sibling cross-products of residuals. This estimator is simply a weighted average, across all families, of the mean (calculated over all sibling pairs) sibling covariance of residuals within each family. Solon, Page, and Duncan discuss four alternative weighting schemes: $\tilde{w}_i = 1$ (scheme 1), $\tilde{w}_i = J_i$ (scheme 2), $\tilde{w}_i = \sqrt{J_i(J_i-1)/2}$ (scheme 3), and $\tilde{w}_i = J_i(J_i-1)/2$ (scheme 4). Weighting scheme 1 gives equal weight to each family, thus ignoring the fact that families with data on more siblings contribute more distinct sibling pairs and hence contain more information. Weighting scheme 4 gives equal weight to each sibling pair, which perhaps weights families contributing many siblings excessively. Weighting schemes 2 and 3 represent compromises between schemes 1 and 4. If desired, one again can assign different weights to families with different survey probabilities simply by multiplying each of the $\tilde{w}_i$ defined above by the inverse of family i's survey probability. This gives eight potential weighting schemes. Finally, given $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_\epsilon^2$, an estimate of the sibling correlation in status, ρ, is simply $\hat{\rho} = \hat{\sigma}_\alpha^2/\hat{\sigma}_\epsilon^2$.

For each of the four different choices of $\tilde{w}_i$ described above, the estimator in (3) can be expressed equivalently as

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^{N} \overline{w}_i \{\sum_{j=1}^{J_i-1} \sum_{j'=j+1}^{J_i} \hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}\}}{\sum_{i=1}^{N} \sum_{j=1}^{J_i-1} \sum_{j'=j+1}^{J_i} \overline{w}_i}, \tag{4}$$

with $\overline{w}_i = 2/J_i(J_i-1)$ for scheme 1, $\overline{w}_i = 2/(J_i-1)$ for scheme 2, $\overline{w}_i = \sqrt{2/J_i(J_i-1)}$ for scheme 3, and $\overline{w}_i = 1$ for scheme 4. As before, differences across families in survey probabilities can be incorporated by multiplying each choice of $\overline{w}_i$ by the inverse of family i's survey probability.[5] The representation in (4) is useful because it implies that each of the alternative estimates $\hat{\rho}$ described above can be obtained simply by running an appropriately weighted regression of $\hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}/\hat{\sigma}_\epsilon^2$, a vector of scaled products of residuals for all distinct sibling pairs in the sample, on a constant. Estimates of $\hat{\rho}$ are presented below.

The analysis so far has assumed that both the variance and the sibling covariance in adult economic status are constant across all families and sibling pairs. As discussed in the introduction, there is little reason to expect this assumption to be true. That is, family characteristics might systematically influence not only mean economic status but also the degree to which siblings resemble each other in their economic status. For example, if wealthy families make less equalizing investments in children's human capital than poorer families do, then the within-family variation in children's educational attainment will be larger in wealthy families and consequently the sibling correlation in (the unexplained component of) educational attainment will be relatively low in wealthy families.[6] More generally, the conjecture that sibling resemblance in status varies systematically across types of families and sibling pairs can be examined empirically by investigating whether the sibling correlation in the unexplained component of economic status depends on family and sibling pair characteristics.

[5]Thus, the four additional weighting schemes are $\overline{w}_i = 2/p_i J_i(J_i-1)$ for scheme 5, $\overline{w}_i = 2/p_i(J_i-1)$ for scheme 6, $\overline{w}_i = \sqrt{2/p_i^2 J_i(J_i-1)}$ for scheme 7, and $\overline{w}_i = 1/p_i$ for scheme 8.

[6]In other words, relatively little of the overall residual variation in educational attainment for children from wealthy families will be between-family variation.

I perform such an analysis in multiple steps. In the first step, estimates of (1) are obtained and residuals, $\hat{\epsilon}_{ij}$, are computed. In contrast to analyses that estimate a common sibling correlation across all families, a large set of background characteristics are included in X because now interest centers on whether the sibling correlation in economic status varies with family and sibling pair characteristics *after* allowing for the effects of family background on mean economic status. In the second step, the squared residuals from the first step, $\hat{\epsilon}_{ij}^2$, are regressed on X_{ij} — thereby allowing the residual variance in status to vary with individual and family background characteristics — and predicted values, $\widehat{\hat{\epsilon}_{ij}^2}$, are constructed from these second step estimates. These predicted values represent an estimate of the variance in economic status for sibling j from family i conditional on the observed background characteristics, X_{ij}. Finally, the third step involves constructing a scaled product of sibling residuals, $\hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}/\sqrt{\widehat{\hat{\epsilon}_{ij}^2}\widehat{\hat{\epsilon}_{ij'}^2}}$, for each distinct sibling pair in the sample and regressing this variable on a set of family background variables and variables measuring the sex composition of the sibling pair and the sibling age difference. The dependent variable in this third step is clearly an observation-level analog of the sibling correlation and the results from the third step regression reveal how the sibling correlation varies with family and sibling pair characteristics. Third step estimates for each of the eight weighting schemes (i.e., choices of $\overline{w}_i$) described earlier are presented below.

3 Data Issues

The empirical analysis uses data from the National Longitudinal Survey of Youth (NLSY), a panel survey of individuals who were between the ages of 14 and 22 at the time of the initial 1979 interview. An important feature of the NLSY is that the primary sampling units were households, with all youths born between 1957 and 1964 and residing in the surveyed households at the time of the 1979 interview becoming respondents. Because of this sampling scheme, the NLSY included 5,863 siblings from 2,448 families at the time

of the first interview.[7] The sibling groups in the NLSY range in size from 2 to 6 siblings, although the majority of families with multiple sibling respondents contribute just a pair of siblings.

In the present paper, I present evidence on how sibling correlations in educational attainment and cognitive ability vary with family and sibling pair characteristics. Cognitive ability is measured by the individual's (regression-adjusted) performance on the Armed Forces Qualifications Test (AFQT), an aptitude test administered to NLSY respondents in 1980. Both educational attainment and AFQT score capture significant dimensions of skill that are clearly related to earnings.[8] In future versions of the paper, I plan to report results from a similar analysis of earnings correlations among NLSY siblings, but these results are not yet available.[9]

The empirical methodology described earlier requires a large sample of families with data on economic status for two or more siblings. My analysis of the sibling correlation in educational attainment uses two complementary samples. The first sample (sample A) consists of all NLSY respondents who report schooling data at age 25 or older and have at least one sibling in the NLSY who also reports schooling data at age 25 or older. Individuals with schooling data only from before age 25 are dropped so that measured schooling likely reflects final educational attainment. For the same reason, the latest reported information on years of schooling is used for the individuals included in the sample.[10]

---

[7]This count only includes respondents who are specifically classified as the "brother" or "sister" of at least one of the other surveyed youths in the household. Pairs of youths who are specifically classified as step-siblings, foster siblings, adopted siblings, or who are coded as belonging to the ambiguous residual categories "natural, step, or adopted siblings" or "siblings or cousins" are not counted as siblings. Unfortunately, the NLSY does not distinguish full biological siblings from half-siblings. Thus, the sample of siblings presumably includes some half-siblings as well as any pairs of step-siblings, foster siblings, adopted siblings, or cousins who were misclassified as siblings.

[8]Neal and Johnson (1996) document the strong relationship between AFQT scores and earnings.

[9]Analyzing sibling correlations in earnings introduces two complicating factors. First, the substantial sex difference in labor force activity implies that performing separate analyses for brother pairs, sister pairs, and brother-sister pairs may be necessary. Second, because earnings are time-varying, an empirical model of earnings must allow for both permanent and transitory error components, thereby making the empirical analysis more complex. Solon, et. al. (1991) provide a clear description of how transitory variation in earnings complicates estimation of the sibling correlation in permanent income.

[10]The analysis uses NLSY data through the 1996 survey round. Respondents who were interviewed in 1996 were 31-39 years old at the time of the interview.

Since all NLSY respondents were born between 1957 and 1964, the maximum possible age difference between sibling respondents is 8 years. Thus, any sample containing only NLSY respondents provides only limited scope for exploring whether the sibling correlation in economic status varies with the sibling age difference. Fortunately, each NLSY respondent was asked to report the age, sex, and educational attainment of each of their siblings in the 1993 survey round. These questions yield data on the educational attainments of the non-NLSY siblings of NLSY respondents, thereby providing data on sibling pairs that are more than 8 years apart in age.

Of course, one might be concerned about the accuracy of a respondent's reports about his *sibling's* age and education, especially if the sibling is much older or younger than the respondent. To minimize these measurement error concerns, I only augment sample A with a select set of non-NLSY siblings of NLSY respondents. Specifically, for a non-NLSY sibling to be included in the expanded sample (sample B), I require that at least two NLSY respondent siblings make reports on the non-NLSY sibling's age, sex, and educational attainment *and* that all of the data on the non-NLSY sibling reported by the NLSY respondent siblings agree exactly. Moreover, consistent with the selection criteria for sample A, I also require that all of the non-NLSY siblings be at least 25 years old. There are 697 non-NLSY siblings of NLSY sibling respondents for whom these sample inclusion criteria are met.

The other sample used in the empirical analysis (sample C) consists of all NLSY respondents who took the AFQT and have at least one sibling in the NLSY who also took the AFQT. The numbers of families, individuals, and sibling pairs in each of the three samples are shown in Table 1.

Table 2 presents weighted means and standard deviations of the status measures and family background characteristics for each of the three samples. Mean educational attainment is about 13.3 years and mean performance on the AFQT was at the 47th percentile.[11]

[11]The AFQT was administered to the sample members in 1980, when they ranged between 15 and 23 years of age.

There is considerably greater age variation in sample B, which contains the 697 siblings of NLSY respondents who were either too old or too young to be included in the NLSY themselves, than in the samples limited only to NLSY respondents.

The remainder of table 2 presents summary statistics for observed family background characteristics. For those characteristics shared in common by all siblings, I assign each sibling in a given family the mean value (or, for indicator variables, the modal value) reported by *all* NLSY-respondent siblings in the family.[12] Because the samples essentially consist only of individuals who have at least one sibling within 7 years of age, average sibship size in the sample is quite large. However, none of the sample means for the other background characteristics appear to be unrepresentative.

4 Empirical Analysis

4.1 Raw Sibling Correlations in Economic Status

Table 3 presents traditional estimates of the sibling correlation in economic status, $\hat{\rho}$, for each of the eight weighting schemes described above.[13] As noted earlier, these estimates impose the assumption that both the variance and sibling covariance in educational attainment and AFQT score do not vary with family or sibling pair characteristics. In the rows labelled "no family background controls", the reported correlations represent the residual sibling correlation in status after controlling for only the respondent's sex, birth year, and age at the time the status measure was obtained. These sibling correlations are comparable to those that have appeared frequently in the literature. In the rows labelled "with family background controls", the reported correlations represent the residual sibling correlation in status after controlling for an extensive set of additional family background characteristics.

[12]For example, if the mean of reported father's education in a group of 3 siblings is 13.5 years, the categorical variable for father's educational attainment between 13 and 15 years takes the value 1 for all 3 siblings and the other categorical variables for father's education take the value 0 for all 3 siblings.

[13]The NLSY only contains individual sampling weights. I create family weights by taking the geometric mean of the individual weights of all NLSY respondent siblings in each family. Over 97% of the variation in individual sampling weights is between families.

A comparison of the sibling correlation estimates with and without family background controls reveals how much of the sibling correlation in status can be explained by observable components of family background.

Without family background controls, the estimated sibling correlations in educational attainment range from .51-.57 in sample A and from .46-.56 in sample B. These magnitudes are very similar to those found in the previous studies. The estimated sibling correlations in AFQT score without family background controls are somewhat larger, ranging from .62-.67, and are also larger than the sibling correlations in IQ reported in the literature (Chamberlain and Griliches 1977, Olneck 1977). That the sibling correlation for AFQT scores is larger than for IQ scores is not surprising, given that the AFQT measures skills that are likely to be influenced by both nature and nurture whereas IQ tests attempt to isolate "innate" intelligence. At the same time, if endowed ability is an important determinant of the AFQT score, the greater sibling correlation in AFQT scores relative to educational attainment is consistent with the notion that parents' investments in children's human capital tend to reinforce sibling differences in endowments.[14]

Controlling for observable dimensions of shared family background reduces the residual sibling correlations in economic status substantially. The nature of the relationship between economic status and the various observable components of family background is summarized in table 4, and will be discussed shortly. For the moment, however, I simply note that the observable part of family background comes nowhere close to fully explaining the sibling correlation in status. In particular, table 3 indicates that the observable dimensions of family background (e.g., sibship size, parents' education, race, etc.) can explain only about

[14]In particular, suppose that endowed ability can be decomposed as $A_{ij} = f_i + s_{ij}$, where f_i is a common family component and s_{ij} is a sibling-specific component. Also, suppose that educational attainment is given by $E_{ij} = \phi f_i + \nu_i + \theta s_{ij} + \omega_{ij}$, where ν_i and ω_{ij} are, respectively, family-specific and individual-specific determinants of educational attainment that are unrelated to ability. Then, the sibling correlation in ability is $\sigma_f^2/(\sigma_f^2 + \sigma_s^2)$ and the sibling correlation in educational attainment is $(\phi^2\sigma_f^2 + \sigma_\nu^2)/(\phi^2\sigma_f^2 + \sigma_\nu^2 + \theta^2\sigma_s^2 + \sigma_\omega^2)$. If more able families tend to invest in more education, then $\phi > 1$. If parents tend to make investments in children's human capital that reinforce innate ability differences, then $\theta > 1$. Clearly, a tendency towards reinforcing investments in children's human capital makes it more likely that the sibling correlation in educational attainment will be smaller than the sibling correlation in ability, as observed in the data.

30-40 percent of the sibling correlation in status. Thus, the factors that account for the majority of the sibling resemblance in economic outcomes are not readily observable.

Finally, it must be noted that the standard errors reported in table 3 are not adjusted for the sampling error that is built into the dependent variable, $\hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}/\hat{\sigma}^2_{\epsilon}$, from the estimation of (1). Consequently, the reported standard errors are likely to be too small.[15] This same remark applies to tables 5-8, which also contain estimation results from regressions in which the dependent variable is itself constructed from previous estimates. I plan to follow Solon, et. al. (1991) and Solon, Page, and Duncan (1997) and estimate consistent standard errors via balanced half-sample replications or some other appropriate resampling method, but I have not yet implemented this procedure. Thus, the reported standard errors should be viewed as suggestive at best.

4.2 Effects of Family Background on Mean Economic Status

The results in table 3 indicate that observable family background explains a substantial minority of the sibling correlation in status. Table 4 illustrates how the economic status measures are related to specific family background characteristics. These estimates represent the first step in the procedure for obtaining estimates of how the sibling correlation in status varies with family and sibling pair characteristics.

On the whole, the results in table 4 are very similar for both measures of status and are quite consistent with the extensive literature on the effects of family background on human capital and earnings. For example, other things equal, educational attainment and AFQT scores are lower for individuals who come from larger sibships, who were raised in non-intact families, and who have less educated parents. Females complete more schooling and have higher AFQT scores than males. Hispanics and blacks both have dramatically lower AFQT scores than observationally equivalent whites but blacks complete more schooling than similar whites. Other than the respondent's age when the status measure was obtained,

[15] However, the reported standard errors do allow for an arbitrary correlation structure between all of the sibling pair observations from a given family.

none of the additional explanatory variables listed in the note to table 4 have a significant impact on economic status. The (unreported) coefficients on the respondent's age indicate that individuals who report education when older have completed more schooling and that individuals who took the AFQT when older earned higher scores.

Table 5 presents least squares estimates of the relationship between $\hat{\epsilon}_{ij}^2$, the squared residuals from the first step estimates of (1), and the detailed demographic and family background measures. These estimates represent the second step in the procedure for obtaining estimates of how the sibling correlation in status varies with family and sibling pair characteristics. The squared education residuals appear to be quite weakly related to most of the observable family background measures. The variance in educational attainment seems to be larger for children of immigrants, which is not surprising given the heterogeneity of the immigrant pool, and perhaps for males. The squared AFQT residuals, on the other hand, are more strongly related to measures of family background. In particular, the variance in AFQT scores is smaller for females, blacks, Hispanics, children from large sibships, and (possibly) children with college-educated parents. The lower variability in AFQT scores among individuals with these characteristics presumably reflects greater homogeneity in opportunities or preferences for individuals belonging to these groups.

4.3 Effects of Family Background on the Sibling Correlation in Status

Tables 6-8 present results illustrating how the sibling correlations in educational attainment and AFQT scores vary with characteristics of the family and sibling pair. These estimates are the main focus of the paper. They are obtained by regressing $\hat{\epsilon}_{ij}\hat{\epsilon}_{ij'}/\sqrt{\widehat{\epsilon_{ij}^2}\widehat{\epsilon_{ij'}^2}}$ on a set of family background and sibling pair characteristics, where $\hat{\epsilon}_{ij}$ is the residual from the first step regression and $\widehat{\epsilon_{ij}^2}$ is the estimated conditional variance of the residuals from the second step regression. Estimates are presented for each of the eight weighting schemes discussed earlier. A very small number of sibling pair observations are lost from the various third step regressions because the predicted values of $\widehat{\epsilon_{ij}^2}$ are negative for a few observations. In

addition, 19 non-twin sibling pair observations for which the reported sibling age difference is less than 7 months are dropped because these sibling pairs are not likely to be full biological siblings.

Although the estimated coefficients are not very precise (even without adjusting for the sampling error introduced by the estimation in the first and second steps), some interesting patterns still can be discerned. Table 6 reports the results for the sibling correlation in educational attainment using data only on sibling respondents in the NLSY (sample A). The estimates uniformly indicate that the residual sibling correlation in educational attainment is much lower among individuals whose fathers completed exactly a high school degree (the excluded category) than among individuals with either more or less educated fathers. Since father's education is the best proxy for family income in these data, this finding suggests that the sibling correlation in educational attainment is higher in both rich and poor families than in middle income families. Moreover, the magnitudes of these effects are very large; the point estimates imply that the residual sibling correlation in educational attainment is around 50 percent higher (.45 versus .30) for children of college educated fathers than for children of high school educated fathers.

The sibling correlation in educational attainment appears to be decreasing in mother's educational attainment. Once again, the point estimates suggest that the changes in the sibling correlation across different categories of maternal education are substantial. In contrast, there is no evidence that the sibling correlation in educational attainment differs across brother pairs, sister pairs, and brother-sister pairs.

Table 6 also illustrates how the sibling correlation in educational attainment varies with the sibling age difference. Recall that all sibling pairs in this sample are separated by less than 8 years, with the vast majority separated by 3 years or less. Each estimated coefficient measures the average difference in the sibling correlation for sibling pairs in the corresponding age difference category relative to siblings who are one year apart in age.[16]

[16]The age difference categories are constructed from sibling age differences measured in months. Thus,

The point estimate on the twin pair indicator is positive and large in some specifications but this result is not robust across weighting schemes. The other notable finding is that the sibling correlation appears to be much *larger* among sibling pairs who are furthest apart in age (i.e., 6 years or more). Thus, in this sample, siblings who are born 6 years apart or more have the most similar educational attainments.

Table 7 shows the results for the sibling correlation in educational attainment using the expanded sample (sample B). The patterns of estimated coefficients on the father's education and mother's education variables are qualitatively similar to those reported in table 6. However, the relationship between the sibling correlation in educational attainment and the sibling age difference is very different. In particular, the sibling correlation falls sharply once the sibling age difference exceeds 8 years. Thus, siblings who are sufficiently far apart in age now appear to have much *less* similar educational outcomes than more closely spaced siblings. In addition, the large positive coefficient on the indicator for a sibling age difference of 6 years or more in table 6 disappears in table 7. The instability of the result in table 6 is not too surprising given that sample A contains only 184 sibling pairs separated by 6 years or more. In contrast, sample B contains 563 sibling pairs separated by 6 or 7 years.

Finally, table 8 presents estimates showing how the sibling correlation in AFQT scores varies with background characteristics. Like the sibling correlation in educational attainment, the sibling correlation in AFQT scores is U-shaped in father's education and the differences in the sibling correlation across the various categories of father's education are large. Somewhat surprisingly, the sibling correlation in AFQT scores appears to increase with maternal education. This contrasts with the evidence in table 6 showing that the sibling correlation in educational attainment *decreases* with maternal education. The estimated coefficients on the sibling age difference variables exhibit no distinct pattern.

---

the category "sibling age difference = 2 years" includes all sibling pairs separated by between 18 and 30 months.

4.4 Discussion of Findings

The empirical analyses in tables 6-8 yield two main results about the relationship between sibling correlations in economic status and family and sibling pair characteristics. First, the sibling correlation in both educational attainment and AFQT scores is U-shaped in father's education. Since father's education is a proxy for the parental family's permanent income in the NLSY data, this finding suggests that siblings from both rich and poor families tend to accumulate more similar stocks of human capital than do siblings from middle income families. The other main finding is that the sibling correlation in educational attainment is lower for siblings who are sufficiently far apart in age than for siblings who are more closely spaced. However, siblings must be fairly widely spaced (i.e., more than 8 years apart) before the reduced sibling resemblance becomes apparent. There is also some evidence that maternal educational attainment influences the sibling correlation in economic status but this last result is not easily interpreted because the nature of the relationship between the sibling correlation in status and mother's education differs across alternative measures of status.

While not a formal test of any particular model of intrafamily allocation, these results do suggest that certain factors ignored by existing models may play an important role in determining how parents allocate human capital investments among their children. For example, the finding that the sibling correlation in educational attainment is U-shaped in father's education is not predicted by either of the standard models of intrafamily allocation. As noted in the introduction, the Becker-Tomes model suggests that, other things equal, children from wealthier families should have less similar human capital stocks than children from poorer families because wealthy parents can make financial transfers to compensate efficient but unequal human capital investments. Behrman, Pollak, and Taubman (1982) show, however, that if parents care about their children's *earnings* rather than their total wealth, then the allocation of human capital investments depends on how parents weight equity versus efficiency, even in wealthy families. Unless one assumes that parental tastes

for child inequality vary systematically with family wealth, this model gives no reason why the sibling correlation in status should depend on family wealth.

However, it is not hard to think of reasonable modifications of these models that can rationalize the evidence presented above. For example, suppose that the marginal return to human capital decreases with the level of human capital but increases with ability. Furthermore, suppose that all children enter parental preferences symmetrically but that parents are not overly averse to unequal child outcomes. Under these (standard) assumptions, parents will tend to invest more in their more able children. Now suppose, in addition, that the marginal return from human capital investment drops discontinuously after some threshold is attained (e.g., a college degree). In this case, parents clearly might prefer to make initial investments primarily in the most able child and to invest primarily in the less able children only after the most able child has reached the threshold education level. If so, the sibling correlation in measures of human capital or earnings will be U-shaped in parental wealth, consistent with pattern observed in the NLSY data.

Alternatively, even if siblings have identical abilities and parents have equal concern for all of their children, parents might invest in their children unequally if there are increasing marginal returns to human capital over some range.[17] If the marginal return from human capital is increasing only over some "intermediate" range, this also would generate the pattern of sibling correlations that are U-shaped in father's education. In particular, families in some middle income range would have insufficient resources to invest efficiently in all of their children and the presence of locally increasing returns would induce them to concentrate the marginal investment in a single child. In contrast, wealthy families would be able to make (equal) efficient investments in all (equally able) children and poor families would make equal (and minimal) investments in all children.[18] In summary, plausible extensions of the standard models of intrafamily allocation can account for the evidence

[17]For example, the gain in lifetime earnings from completing college might be more than twice as large as the gain from completing only the first two years of college.

[18]Dahan and Gaviria (1998) develop a formal model with these features.

presented above. However, this same evidence suggests that the stronger assumptions (e.g., homotheticity of the earnings production function) sometimes invoked to identify parental preference parameters may not be appropriate.

Next, consider the evidence that the sibling correlation in educational attainment drops off for siblings who are sufficiently far apart in age. The most straightforward explanation for this finding is that closely spaced siblings experience more similar family environments than do widely spaced siblings and that this greater similarity in environment leads to greater similarity in human capital acquisition. Of course, many components of the family environment can change over time — including family income, parents' time allocations, neighborhood quality, and so on — so this result does not identify which of the many time-varying components of family environment play important roles in determining eventual economic status. However, if parents have equal concern for all of their children, one would expect them to try to smooth out secular changes and transitory fluctuations in family environment.[19] Thus, the evidence that siblings born far apart have less similar educational outcomes is consistent with the notion that families find it difficult to smooth out shocks to family environment.

Unfortunately, while the evidence is consistent with this view, it is not totally conclusive because the NLSY does not distinguish half-siblings from full-siblings. Half-siblings share only one biological parent and hence are less similar genetically than full siblings. Consequently, one would expect the correlation in economic status to be lower for half-siblings than for full-siblings. Moreover, half-siblings probably tend to be more widely spaced than full-siblings. Thus, the finding that the sibling correlation in educational attainment is lower for widely spaced siblings might simply result from a disproportionate representation of half-siblings among all pairs of widely spaced siblings. Evaluating the importance of this hypothesis will require data on the frequency and spacing of half-siblings and full-siblings

[19]For example, families that anticipated rising income over time would wish to borrow to give their earlier-born children the same resources as later-borns.

in the population. I hope to provide such evidence in future versions of the paper. Finally, endogenous spacing of births further complicates the interpretation of the evidence on the "effect" the sibling age difference. For example, if parents who are less averse to within-family inequality tend to space births more widely, the observed pattern of lower sibling correlations among more widely spaced siblings could simply reflect unobserved parental preferences.

5 Conclusion

This paper has provided detailed evidence on the sibling correlations in educational attainment and cognitive ability (as measured by AFQT scores). Like many other authors, I find that the sibling correlations in these variables are very large and that observable family background variables can explain less than half of the observed correlations. Furthermore, the sibling correlation in educational attainment is less than the sibling correlation in AFQT scores which at least suggests that, on average, parental investments in children's human capital tends to accentuate differences in endowed ability.

The novel contribution of the paper is its investigation of whether and how these sibling correlations vary with observed family and sibling pair characteristics. Although the resulting estimates are not very precise, the point estimates suggest that the sibling correlation in status varies substantially with certain characteristics. Two findings stand out in particular. First, the correlation in status appears to be lower among siblings with high-school educated fathers than among siblings with either more or less educated fathers. In other words, a larger fraction of the unexplained variation in economic status is within-family variation for individuals whose fathers have exactly a high school degree. This result can be explained by a model of intrafamily allocation in which parents are not too averse to inequality among their children and in which the marginal return to human capital either drops discontinuously or is increasing over certain ranges. Interestingly, under these conditions some families will allocate resources in a way that generates substantially unequal

child outcomes.

The second major finding is that the sibling correlation in educational attainment appears to be considerably lower among widely spaced siblings. This result is consistent with the view that time-varying components of family environment, such as the time paths of family income and parental time allocations, have nontrivial impacts on a child's eventual economic status. Moreover, this result suggests that families are unable to fully smooth out transitory fluctuations and secular trends in the household environment. This limited ability to smooth environmental shocks further constrains the extent to which families can produce equal outcomes among their children.

In summary, the degree of sibling resemblance in economic status appears to vary across both differentially endowed families and differentially spaced sibling pairs. This evidence strongly suggests that differences in family budget constraints, both between families and within families over time, cause systematic differences in how families allocate resources among children. This finding has numerous potential policy implications. For example, it implies that redistributive policies will influence inequality not only by changing the distribution of resources between families but also by changing the allocation of resources within families. Gaining a more detailed understanding of how family allocation decisions evolve over time and vary with family circumstances is an important task for future research.

References

Altonji, Joseph and Thomas Dunn. "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives", in *Research in Labor Economics* vol. 12, Ronald Ehrenberg (ed.). Greenwich, CT: JAI Press, 1991, 269-310.

Becker, Gary. *A Treatise on the Family.* Cambridge, MA: Harvard University Press, 1991 (enlarged edition).

Becker, Gary and Nigel Tomes. "Child Endowments and the Quantity and Quality of Children", *Journal of Political Economy* 84, August 1976, S143-S162.

Behrman, Jere. "Intrahousehold Distribution and the Family", in *Handbook of Population and Family Economics* vol. 1A, Mark Rosenzweig and Oded Stark (eds.). Amsterdam: North-Holland, 1997, 125-187.

Behrman, Jere and Paul Taubman. "Is Schooling 'Mostly in the Genes'? Nature-Nurture Decomposition Using Data on Relatives", *Journal of Political Economy* 97, December 1989, 1425-1446.

Behrman, Jere; Robert Pollak; and Paul Taubman. "Parental Preferences and Provision for Progeny" *Journal of Political Economy* 90, February 1982, 52-73.

Behrman, Jere; Robert Pollak; and Paul Taubman. "Do Parents Favor Boys?" *International Economic Review* 27, February 1986, 31-52.

Behrman, Jere; Mark Rosenzweig; and Paul Taubman. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: the Twins Experiment", *Journal of Political Economy* 102, December 1994, 1131-1174.

Behrman, Jere and Paul Taubman. "Birth Order, Schooling, and Earnings", *Journal of Labor Economics* 4, 1986, S121-S145.

Behrman, Jere; Paul Taubman; and Terence Wales. "Controlling for and Measuring the Effects of Genetics and Family Environment in Equations for Schooling and Labor Market Success", in *Kinometrics: Determinants of Socioeconomic Success within and between Families*, Paul Taubman (ed.). Amsterdam: North-Holland, 1977, 35-96.

Chamberlain, Gary and Zvi Griliches. "More on Brothers", in *Kinometrics: Determinants of Socioeconomic Success within and between Families*, Paul Taubman (ed.). Amsterdam: North-Holland, 1977, 97-124.

Dahan, Momi and Alejandro Gaviria. "Parental Actions and Siblings' Inequality", Inter-American Development Bank OCE Working Paper no. 389, December 1998.

Griliches, Zvi. "Sibling Models and Data in Economics: Beginnings of a Survey", *Journal of Political Economy* 87, October 1979, S37-S64.

Neal, Derek and William Johnson. "The Role of Premarket Factors in Black-White Wage Differences", *Journal of Political Economy* 104, October 1996, 869-895.

Olneck, Michael. "On the Use of Sibling Data to Estimate the Effects of Family Background, Cognitive Skills, and Schooling: Results from the Kalamazoo Brothers Study", in *Kinometrics: Determinants of Socioeconomic Success within and between Families*, Paul Taubman (ed.). Amsterdam: North-Holland, 1977, 125-167.

Solon, Gary. "Intergenerational Mobility in the Labor Market", 1998, mimeo.

Solon, Gary; Mary Corcoran; Roger Gordon; and Deborah Laren. "A Longitudinal Analysis of Sibling Correlations in Economic Status", *Journal of Human Resources* 26, Summer 1991, 509-534.

Solon, Gary; Marianne Page; and Greg Duncan. "Correlations between Neighboring Children in their Subsequent Educational Attainment", 1997, mimeo.

Table 1: Composition of samples: families, individuals, and sibling pairs

Sample	Status measure	Size of sibling group	Number of families	Number of individuals	Number of sibling pairs
A	Education	2	1636	3272	1636
		3	543	1629	1629
		4	138	552	828
		5	24	120	240
		6	3	18	45
		Total	2344	5591	4378
B	Education	2	1262	2524	1262
		3	724	2172	2172
		4	244	976	1464
		5	81	405	810
		6	25	150	375
		7	5	35	105
		8	2	16	56
		10	1	10	45
		Total	2344	6288	6289
C	AFQT score	2	1598	3196	1598
		3	534	1602	1602
		4	137	548	822
		5	23	115	230
		6	3	18	45
		Total	2295	5479	4297

Sample A consists of all groups of 2 or more NLSY sibling respondents with data on educational attainment at age 25 or older. Sample B consists of all of the individuals in sample A plus 697 of their non-NLSY siblings. Information on age, sex, and educational attainment of all non-NLSY siblings of NLSY respondents was obtained in the 1993 survey round. These 697 individuals represent the complete set of non-NLSY siblings for whom (1) at least two NLSY respondent siblings provided information on age, sex, and educational attainment and (2) the multiple reports by the NLSY respondent siblings matched exactly. Sample C consists of all groups of 2 or more NLSY sibling respondents with AFQT score data.

Table 2: Summary statistics for the various samples of individuals

	Sample A		Sample B		Sample C	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.
Years of education	13.31	2.52	13.31	2.49		
AFQT score (percentile)					47.00	29.09
Age when status measured	34.51	2.60	34.72	3.48	19.21	2.07
Sibship size	4.71	2.27	4.84	2.30	4.70	2.25
First born	.186		.206		.185	
Second born	.282		.272		.281	
Third born	.196		.196		.196	
Fourth born	.133		.129		.133	
Fifth born	.081		.078		.081	
Mother's age at child's birth	25.74	5.67	25.62	5.72	25.78	5.72
Father's age at child's birth	29.05	6.68	28.83	6.88	29.12	6.75
Mother's education < 12 years	.336		.343		.337	
Mother's education 13-15 years	.150		.147		.150	
Mother's education ≥ 16 years	.101		.099		.101	
Father's education < 12 years	.355		.356		.352	
Father's education 13-15 years	.154		.155		.154	
Father's education ≥ 16 years	.182		.177		.181	
≥ 1 parent is immigrant	.096		.092		.094	
≥ 1 sibling lived in non-intact family at age 14	.254		.249		.252	
Female	.484		.485		.485	
Black	.154		.155		.156	
Hispanic	.061		.059		.059	

The summary statistics are weighted using NLSY sample weights. For family background characteristics shared in common by all siblings in a family, the mean value (modal value for indicator variables) reported by all NLSY-respondent siblings in the family is assigned to each sibling.

Table 3: Estimates of the sibling correlation in economic status for alternative samples, measures of status, and weighting schemes

	Weighting scheme							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Education (Sample A)								
No family background controls	.550 (.025)	.536 (.024)	.531 (.024)	.506 (.026)	.570 (.028)	.557 (.027)	.552 (.027)	.529 (.028)
With family background controls	.377 (.024)	.367 (.023)	.363 (.023)	.347 (.026)	.369 (.029)	.353 (.026)	.347 (.026)	.319 (.025)
Education (Sample B)								
No family background controls	.543 (.025)	.514 (.023)	.505 (.022)	.461 (.022)	.558 (.028)	.529 (.025)	.520 (.025)	.477 (.025)
With family background controls	.363 (.024)	.340 (.022)	.333 (.022)	.311 (.025)	.350 (.028)	.318 (.025)	.308 (.024)	.278 (.026)
AFQT score (Sample C)								
No family background controls	.658 (.021)	.661 (.021)	.661 (.021)	.665 (.023)	.619 (.024)	.629 (.024)	.632 (.024)	.655 (.028)
With family background controls	.433 (.023)	.420 (.022)	.416 (.022)	.392 (.023)	.414 (.028)	.399 (.026)	.393 (.026)	.369 (.027)

The different weighting schemes are described in the text. In the rows labelled "no family background controls", the reported estimate represents the sibling correlation in residuals from a regression of the status measure on a female dummy, a second order polynomial in the respondent's age at the time the status measure was obtained, and (in samples A and B) birth year dummies. In the rows labelled "with family background controls", the reported estimate represents the sibling correlation in residuals from a regression of the status measure on the full set of individual and family background variables listed in table 4 and the note following table 4. The reported standard errors allow for arbitrary correlation of the sibling cross-products of residuals within families, but do not adjust for the sampling error in the first stage estimates.

Table 4: Selected coefficient estimates for model of the conditional mean of individual economic status

	Dependent variable and sample					
	Education (Sample A)		Education (Sample B)		AFQT score (Sample C)	
	(1)	(2)	(3)	(4)	(5)	(6)
Sibship size	−.1452	−.1388	−.1315	−.1270	−1.093	−.8885
	(.0206)	(.0248)	(.0197)	(.0237)	(.1983)	(.2787)
Mother's education < 12	−.6802	−.7187	−.6639	−.6939	−7.066	−6.738
	(.0894)	(.1073)	(.0843)	(.1016)	(1.002)	(1.297)
Mother's education 13-15	.5002	.4811	.4956	.4841	4.885	3.726
	(.1163)	(.1314)	(.1119)	(.1256)	(1.405)	(1.583)
Mother's education ≥ 16	1.202	1.156	1.287	1.262	9.878	8.255
	(.1727)	(.1933)	(.1637)	(.1819)	(1.875)	(1.938)
Father's education < 12	−.4781	−.5111	−.3944	−.4051	−7.488	−7.201
	(.0894)	(.1085)	(.0850)	(.1034)	(1.041)	(1.343)
Father's education 13-15	.5258	.5405	.5173	.5318	3.565	3.088
	(.1212)	(.1410)	(.1133)	(.1312)	(1.394)	(1.557)
Father's education ≥ 16	1.530	1.574	1.533	1.581	13.471	13.753
	(.1499)	(.1620)	(.1431)	(.1536)	(1.664)	(1.717)
≥ 1 immigrant parent	.2175	.0941	.2214	.1251	−.4824	−2.783
	(.1322)	(.1643)	(.1261)	(.1553)	(1.349)	(1.867)
≥ 1 sibling lived in non-intact family at age 14	−.3183	−.2647	−.2771	−.2347	−3.063	−2.356
	(.0825)	(.1003)	(.0813)	(.1017)	(.8675)	(1.229)
Female	.4999	.3890	.4520	.3403	1.889	2.338
	(.0567)	(.0674)	(.0535)	(.0639)	(.5815)	(.7703)
Black	.3755	.3136	.3653	.2959	−18.241	−21.130
	(.0899)	(.0956)	(.0865)	(.0943)	(.9164)	(1.066)
Hispanic	−.0851	−.1081	−.1499	−.1811	−9.370	−10.881
	(.1125)	(.1336)	(.1064)	(.1261)	(1.270)	(1.550)
Sample weights?	No	Yes	No	Yes	No	Yes
R^2	.304	.335	.297	.326	.434	.408
Sample size	5591	5591	6288	6288	5479	5479

Additional explanatory variables are a second order polynomial in the respondent's age when the status measure was obtained, dummies for year of birth (samples A and B only), a second order polynomial in the mother's age at the time of the respondent's birth, a second order polynomial in the father's age at the time of the respondent's birth, dummies for birth order, and indicators for missing data on mother's age, father's age, mother's education, father's education, birth order, and parent's immigrant status. The estimated standard errors allow for arbitrary correlation of the errors among siblings from the same family.

Table 5: Selected coefficient estimates for model of the conditional variance of individual economic status

	Dependent variable and sample					
	Education (Sample A)		Education (Sample B)		AFQT score (Sample C)	
	(1)	(2)	(3)	(4)	(5)	(6)
Sibship size	.1227	−.0711	.0925	−.0828	−12.578	−14.791
	(.0801)	(.0796)	(.0728)	(.0747)	(5.269)	(6.865)
Mother's education < 12	−.2588	−.4337	−.3110	−.4860	−37.397	−35.363
	(.3192)	(.3584)	(.2953)	(.3355)	(24.169)	(30.596)
Mother's education 13-15	−.3306	−.3847	−.3252	−.4058	−18.489	−16.896
	(.3628)	(.3466)	(.3382)	(.3244)	(32.803)	(33.873)
Mother's education ≥ 16	−.1026	−.1794	−.1347	−.3743	−76.293	−59.421
	(.5314)	(.5523)	(.4923)	(.4988)	(49.295)	(49.562)
Father's education < 12	.1099	.0747	.2004	.2969	−25.043	63.772
	(.3237)	(.3698)	(.3012)	(.3443)	(24.449)	(30.960)
Father's education 13-15	.1912	.3557	.2047	.4581	17.471	1.404
	(.3802)	(.4195)	(.3468)	(.3828)	(35.078)	(36.681)
Father's education ≥ 16	.2062	−.0077	.4148	.3131	−48.876	−95.040
	(.4639)	(.4549)	(.4255)	(.4142)	(42.073)	(38.619)
≥ 1 immigrant parent	2.076	1.376	2.190	1.373	45.350	58.710
	(.4686)	(.4827)	(.4498)	(.4612)	(33.770)	(39.903)
≥ 1 sibling lived in non-intact family at age 14	.2263	.4208	.2017	.4558	4.216	20.578
	(.3046)	(.3409)	(.2909)	(.3532)	(22.131)	(29.354)
Female	−.3728	−.5066	−.3112	−.4979	−75.469	−85.313
	(.2160)	(.2295)	(.2022)	(.2131)	(16.530)	(20.701)
Black	−.2755	−.0096	−.3074	−.0636	−230.692	−264.637
	(.3086)	(.3055)	(.2923)	(.3091)	(23.824)	(27.985)
Hispanic	−.1498	.5506	−.2276	.5001	−139.536	−169.165
	(.4001)	(.4606)	(.3778)	(.4434)	(31.881)	(36.481)
Sample weights?	No	Yes	No	Yes	No	Yes
R^2	.017	.012	.018	.013	.056	.046
Sample size	5591	5591	6288	6288	5479	5479

Additional explanatory variables are the same as those listed in the note to table 4. The estimated standard errors allow for arbitrary correlation of disturbances among siblings from the same family but are not adjusted to account for the sampling error built into the dependent variable from the first stage estimates.

Table 6: Selected coefficient estimates for model of the sibling correlation in educational attainment (using sample A)

	Weighting scheme							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Brother pair	.0517	.0437	.0407	.0446	.0357	.0280	.0250	.0231
	(.0542)	(.0482)	(.0465)	(.0450)	(.0644)	(.0573)	(.0550)	(.0478)
Sister pair	−.0310	−.0244	−.0220	−.0049	−.0226	−.0257	−.0269	−.0240
	(.0531)	(.0486)	(.0474)	(.0463)	(.0699)	(.0630)	(.0609)	(.0551)
Sib age diff = 0 yrs	.0602	.1044	.1199	.1728	.0140	.0372	.0449	.0722
(Twin pair)	(.1383)	(.1412)	(.1437)	(.1555)	(.1256)	(.1293)	(.1317)	(.1390)
Sib age diff = 2 yrs	.0765	.0634	.0587	.0388	.0662	.0565	.0529	.0349
	(.0576)	(.0527)	(.0513)	(.0475)	(.0709)	(.0653)	(.0636)	(.0589)
Sib age diff = 3 yrs	.0313	.0228	.0197	.0091	.0004	−.0073	−.0101	−.0227
	(.0646)	(.0569)	(.0545)	(.0471)	(.0805)	(.0717)	(.0687)	(.0594)
Sib age diff = 4 yrs	−.0198	−.0223	−.0232	−.0274	−.0137	−.0174	−.0187	−.0232
	(.0756)	(.0670)	(.0646)	(.0574)	(.0964)	(.0854)	(.0818)	(.0700)
Sib age diff = 5 yrs	.1106	.0758	.0647	.0252	.1090	.0847	.0767	.0520
	(.1132)	(.0953)	(.0901)	(.0742)	(.1076)	(.0935)	(.0894)	(.0778)
Sib age diff ≥ 6 yrs	.2153	.1799	.1688	.1162	.3749	.3174	.2985	.2176
	(.1020)	(.0896)	(.0865)	(.0785)	(.1552)	(.1333)	(.1266)	(.1029)
Mother's educ < 12	.0901	.0857	.0842	.0647	.1448	.1241	.1164	.0739
	(.0610)	(.0563)	(.0553)	(.0581)	(.0808)	(.0740)	(.0719)	(.0662)
Mother's educ 13-15	−.1529	−.1492	−.1480	−.1487	−.1887	−.1893	−.1896	−.1950
	(.0758)	(.0727)	(.0726)	(.0794)	(.0838)	(.0773)	(.0757)	(.0762)
Mother's educ ≥ 16	−.1289	−.1362	−.1389	−.1657	−.1200	−.1214	−.1216	−.1399
	(.1289)	(.1250)	(.1247)	(.1336)	(.1289)	(.1272)	(.1279)	(.1396)
Father's educ < 12	.0919	.0944	.0952	.1053	.1111	.1117	.1118	.1170
	(.0575)	(.0540)	(.0532)	(.0534)	(.0759)	(.0709)	(.0697)	(.0669)
Father's educ 13-15	.1434	.1256	.1188	.0908	.2025	.1745	.1636	.1220
	(.0890)	(.0809)	(.0786)	(.0727)	(.0969)	(.0877)	(.0847)	(.0751)
Father's educ ≥ 16	.1473	.1689	.1768	.2215	.1689	.1759	.1783	.2048
	(.0987)	(.0968)	(.0976)	(.1141)	(.1070)	(.1012)	(.1003)	(.1080)
R^2	.011	.010	.010	.010	.014	.013	.012	.012
Sample size	4358	4358	4358	4358	4358	4358	4358	4358

The different weighting schemes are described in the text. Additional explanatory variables are sibship size, indicators for race (black and Hispanic), an indicator for parents' immigrant status, an indicator for whether at least one sibling was raised in a non-intact family, and indicators for missing data on mother's education, father's education, and parent's immigrant status. The estimated standard errors allow for arbitrary correlation of disturbances among siblings from the same family but are not adjusted to account for the sampling error built into the dependent variable from the 1st and 2nd stage estimates.