Title
Estimating the Effects of Family Background on the Return to Schooling

Author
Deschenes, Olivier

Publication Date
2002-03-01

ESTIMATING THE EFFECTS OF FAMILY BACKGROUND ON THE RETURN TO SCHOOLING.

Olivier Deschênes[1]
Department of Economics
University of California, Santa Barbara

June 2001
This version: March 2002

Abstract: This paper examines the causal association between family background characteristics--parental education and family size--and returns to schooling using data from the Occupational Change in a Generation Survey. I first develop a formal model of schooling and earnings, with heterogeneous returns to education. Family environment is shown to influence the marginal return to schooling through its effects on the marginal benefit and the marginal cost of an additional year of education. Using two types of exclusion restrictions, I find that men raised in larger families have substantially lower returns to education, while the combined effects of parental education on the returns to education are more modest. I also examine the difference between OLS and TSLS estimates of the return to schooling. Like other "supply-side" IV studies of the causal effect of education, this paper documents TSLS estimates that are larger than the corresponding OLS estimates. The results of this paper provide an alternative explanation for this phenomenon: constant marginal returns to schooling, combined with a negative ability bias and a positive self-selection bias (i.e. non-hierarchical sorting).

JEL classification: J24, I21, C21.
Keywords: Return to education, Self-selection, Ability bias, Family Background.

[1] email: olivier@econ.ucsb.edu, web: http://www.econ.ucsb.edu/~olivier
I am grateful to Orley Ashenfelter and David Card for many helpful discussions. I also thank Theodore Bergstrom, Ken Chay, David Lee, Peter Kuhn, and seminar participants at the University of California at Berkeley and Santa Barbara for their comments.

1. Introduction

Economists have long been interested in the effects of family environment on the subsequent labor market success of individuals.[1] Part of this interest stems from the strong correlation between the educational attainment of parents and children, which may contribute to the transmission of socioeconomic status and inequality across generations. In recent years this attention has been heightened by a major transformation of the American family, and by the increasing role of education as a determinant of economic well-being.[2]

Recent studies of the causal association between schooling and earnings have emphasized the heterogeneity in the economic return to an additional year of education across otherwise comparable individuals.[3] Despite increased attention to the possibility of heterogeneous returns to education across individuals, there is still considerable uncertainty about the mechanism generating this heterogeneity. Part of this uncertainty is attributable to the absence of a formal model that explicitly recognizes the possibility that the causal return to schooling varies with observable characteristics, like family background variables.

This paper examines the relationship between family background characteristics and the return to schooling subsequently received by individuals in the labor market. The paper begins by documenting several features of the relationship between family background factors, educational attainment and earnings. Using a large sample from the 1973 Occupational Change in a Generation Survey, I find that men raised by better-educated parents acquire more schooling and have higher earnings, while those raised in larger families are less educated and have lower earnings. Next, I show that the negative effect of family size varies with the gender composition of the sibship. In particular, holding family size and background constant, I find that men raised with more sisters have substantially lower schooling and earnings. These patterns are robust to

[1] See for example Becker (1964), Taubman (1977) and Griliches (1979).

[2] See for example Haveman and Wolfe (1993), Mayer (1997) and Danziger and Waldfogel (2000).

[3] See the evidence contained in Card and Krueger (1992), Heckman, Layne-Farrar and Todd (1996), Altonji and Dunn (1996) and Ashenfelter and Rouse (1998). Heckman and Vytlacil (1998) and Card (1999) discuss theoretical models of heterogeneous returns to education.

a wide variety of specifications.

The contribution of this paper is to develop and implement a formal model of schooling and earnings to interpret these patterns. In light of the recent instrumental variables studies of the causal effect of education, the return to schooling is allowed to vary across individuals, and in particular with the observable characteristics of the family. This distinguishes the current paper from most of the literature, which typically assumes that the return to schooling is constant across the population or is a single random variable. A key implication of the model is that family background can potentially affect both the payoff to an additional year spent in school <u>and</u> the level of acquired schooling. Therefore, a complete assessment of the link between family background and the return to schooling must examine the effect family background on both the marginal benefit and the marginal cost of schooling.[4]

An extensive literature has clearly established that the identification of the causal relationship between schooling and earnings requires an exogenous source of variation in educational choices. It follows naturally that the identification of the parameters describing the costs and benefits of schooling requires two types of exclusion restrictions. The identification of the parameters in the marginal benefit function requires the existence of an observable variable affecting schooling choices only through its effect on the cost of schooling (i.e. an instrumental variable for schooling). Similarly, the identification of the parameters from the marginal cost function requires the existence of an observable variable affecting schooling choices only through its effect on the benefit to schooling. In this study, measures of school quality are used as variables that affect the benefit, but not the cost of schooling, conditional on family background characteristics. Then, following Butcher and Case (1994), I exploit the randomness embodied in the gender composition among siblings holding family size constant, as a variable influencing only the cost of schooling. These two exclusion restrictions allow the estimation of the average causal effect of education, and of the parameters describing the effect of family background on the return to schooling. Moreover, since the effects of gender composition on educational attainment are presumably larger

[4] This possibility was overlooked by Altonji and Dunn (1996) and Ashenfelter and Rouse (1998) who analyzed interactions of schooling and parental education in earnings regressions.

for poorer families (conditional on family size), it is possible to test the assumption that sibling gender composition has an independent effect on earnings. The results of a series of specification tests provide no evidence against the hypothesis that conditional on family size, sibling gender composition is an exogenous determinant of schooling.

The results can be summarized as follows. First, men raised in larger families have significantly lower returns to education.[5] This finding is entirely attributable to the lower benefits per year of education received by individuals raised with more siblings (i.e. it is not related to differences in the costs of schooling). The *combined* effects of parental education on the returns to schooling are more modest. Men who were brought up by better-educated fathers have higher marginal returns to schooling, while those with better-educated mothers have lower marginal returns to schooling. In other words, own education and father's education are q-complements in the production of earnings capacity, while own education and mother's education are q-substitutes. In addition to their opposite signs, the analysis suggest that these effects of parental education operate through distinct mechanisms relating familial environment and returns to education. Father's education is associated with higher benefits per year of education, while mother's education is associated with lower costs per year of education, consequently raising education levels but lowering the marginal return.

The paper is organized as follows. Section 2 describes the data and provides a preliminary descriptive analysis. Section 3 presents a model of schooling and earnings emphasizing the contribution of family background characteristics to the heterogeneity in the returns to schooling across the population. Section 4 presents the identification and estimation strategies used in this paper. Section 5 discusses the empirical results. Section 6 examines the robustness of the findings, with a special attention to the validity of the exclusion restrictions used and the sensitivity of the estimates to measurement error. Section 7 concludes.

[5] Blake (1989) documents similar patterns in the relationship between number of siblings and various aptitude test scores.

2. Preliminary Analysis and Data Description

An ideal data set for the study of the effects of family background on the return to schooling would provide detailed information on current labor market outcomes of individuals, as well as information on the characteristics of their families during the childhood years. The data in this paper are taken from the 1973 Occupational Changes in a Generation survey (OCG). This data set provides a unique source of family background information (number of siblings, education of both parents, family income at age 16, state of birth, etc.)[6] While other data sets like the NLSY and the PSID contain similar family background information, their small sample sizes, missing data problems, and non-representativeness limits the interpretation of any result derived from them. The data from the Occupational Changes in a Generation survey are drawn from a large and representative sample of the adult male population in the United States. The survey was carried out as an eight page mailout-mailback supplemental questionnaire to individuals in the sampling frame for the 1973 March Current Population Survey (CPS). The target population consisted of civilian male aged 20 to 65. There were 37,694 respondents in the survey.[7]

This study focuses on a sample of men aged 24-65, born in the U.S., and excluding Hispanics.[8] Table 1 presents some summary statistics for the baseline sample and the subsample of prime-aged workers. All statistics reported in this paper are weighted by the OCG sample weights. As column 1 shows, the OCG sample provides a nationally representative sample of the population of men aged 20-65. The family background information reveals that the average sibship size was about 4 for these cohorts of men. Individuals were also asked to report the education of their parents and the entries in Table 1 show that mothers are slightly better-educated than fathers. A small fraction of individuals did not report their parent's education. Observations with missing data on parental education were imputed with the predicted values from separate regressions of

[6] The OCG was originally designed to help determining the patterns of intergenerational occupation mobility in the United States. See Featherman and Hauser (1978).

[7] Two supplemental samples of black and Hispanic household heads were drawn. The present analysis is based on the sample of 32,986 males aged 20-65 from the March CPS population.

[8] The age restriction was imposed to ensure that individuals have completed their educational investments at the time of the survey. Hispanics were excluded because they were oversampled from the CPS sampling frame.

father and mother's education on other measures of family background and children's education and earnings.[9] The figures in parentheses below each measure of parental education correspond to fraction of observations imputed. The statistical models reported below always include dummy variables indicating whether either parent's education was imputed. Column 2 reports the characteristics of individuals aged 24-65 who earned more than $60 per week on average and worked full-time in 1972.[10] The analyses in this paper will be performed on this subsample of 17,300 observations in column 2. Comparisons of the mean characteristics in columns 1 and 2 indicate no important differences between the baseline sample and the subsample of workers.

The data from the OCG samples is supplemented by characteristics of public schools in each state for the years 1918-1968. In particular, semiannual data from the Biennial Survey of Education covering the years 1918-1958, and annual data from the Digest of Education Statistics starting in 1960 provide information about statewide enrollment, number of teachers, teacher salary and term length. These data have been used in previous studies, notably Card and Krueger (1992), from whom I draw the samples used to supplement the OCG data. Based on state and year of birth, I assigned the average elementary and secondary school quality that was potentially available to each individual if he would have completed the first 12 years of schooling.[11] Focussing on "potential" school quality rather than actual leaves the endogeneity of educational attainment with school quality aside. In this study, I focus two measures of school quality: the pupil-teacher ratio and the relative teacher salary.[12] As Card and Krueger (1992) and Heckman et al. (1996) documented, other measures like term length are only weakly related with returns to schooling and do not vary as much across cohorts. Moreover, using within-family contrasts, Altonji and Dunn (1996) showed that teacher salary and per-pupil expenditure have a substantial effect on wages.

[9] The regressions include controls for the children's years of education, earnings, race, region of residence in 1973 and family background controls like father (or mother's) education, number of siblings and farm residence at age 16.

[10] This corresponds to the weekly earnings of individuals working 40 hours per week at the 1973 federal minimum wage. Weekly earnings were computed by dividing annual earnings from wages and salaries in 1972 by the number of weeks worked in 1972. Individuals born in Alaska and Hawaii were also excluded from the analysis.

[11] That is, each individual is assigned the average school quality in his state of birth when he was aged 6-16.

[12] The average salary of teachers was normalized by the level average wages in each state.

The analysis begins by examining regressions of educational attainment and log earnings on measures of family background and school quality. All models include cohort dummies (for men born between 1910-1919, 1920-1929, 1930-1939, and 1940-1949), a race indicator, 3 region of birth dummies, 3 dummies for the region of residence in 1973 and an indicator for residence in a metropolitan area (SMSA) in 1973. Table 2 presents a variety of reduced-form regressions for years of completed education. The models in columns (1)-(3) add an increasing set of family background characteristics: parental education (column 1), and number of siblings (column 2). Column 3 breaks down the number of siblings variable into number of brothers and sisters.[13] Column (4) adds the 2 measures of school quality. As shown by the F-statistics in row 8 and the t-statistics, the family background characteristics are always individually and jointly significant at the 5% level.

Row 3 shows that men with better-educated parents complete more years of education, while those from larger families acquire less schooling.[14] Rows 4 and 5 confirm this, but also show that the effect of the number of siblings varies with the gender composition of the sibship. For example, the entries in columns (3) and (4) indicate that holding the number of siblings and family characteristics constant, men with more sisters have less schooling, each sister reducing years of education by about 0.05 years (p-value=0.04).[15] School quality, as measured by pupil-teacher ratio and relative teacher pay is strongly correlated with years of education, as indicated by an F-statistic of 75.45 (p-value=0.00). In columns (5)-(8) the models are estimated separately for 4 different birth cohorts (for men born between 1910-1919, 1920-1929, 1930-1939, and 1940-1949). The within-cohort analysis is motivated by the important changes in family structure and school

[13] All models include other measures of the family structure at age 16: a dummy indicating if the respondent lived with both parents at age 16, and a dummy indicating if the respondent lived on a farm at age 16.

[14] This pattern has been documented by others. See for example Butcher and Case (1994) and Card and Lemieux (2001) on the relationship between education attainment and parental education. Blake (1989) provides an extensive analysis of family size and educational attainment.

[15] This finding of a negative and significant effect of sibship gender composition on the education levels of men is contrary to the finding of Butcher and Case (1994) who analyzed the 1985 wave of the PSID and the NLSW. They documented that holding family size constant, women with more sisters have less schooling, but that the educational attainment of men is unrelated to the gender composition of the sibship. Using data on younger cohorts from the NLSY, Kaestner (1997) concluded that siblings sex composition has little effect on the educational attainment of young adults. Using data from the OCG, SIPP and NSFH, Hauser and Kuo (1998) found little evidence of any gender composition effect on the education of women born 1910-1964.

quality for the cohorts born between 1910 and 1949.[16] Those important changes might not be fully captured by the cohort dummies included in the models (1)-(4). Again, all of the family background characteristics and school quality measures (except one) are individually significant at the 5% level. As indicated by the F-statistics in rows 8 and 9, family background and school quality are jointly significant for each of the 4 birth cohorts. Column (9) provides the F-statistics for testing the equality of the effects across the 4 cohorts. The results indicate that the effects of parental education on schooling are similar across cohorts (p-values=0.06). The effect of gender composition is more variable across cohorts, especially the effect of number of brothers. For the cohorts of men born after 1920, a higher number of sisters is associated with lower schooling, given family size, with an especially strong effect for the men born in the 1940s (-0.12, with a p-value of 0.01). Finally, the effects of school quality on acquired schooling are similar across the 4 cohorts.

Table 3 presents reduced-form regressions (i.e. excluding schooling) of log earnings on the same specifications of family background and school quality displayed in Table 2. The estimated coefficients have the same signs as those reported in Table 2, but they are smaller in magnitude. In the estimated models, all but one of the family background variables are individually significant at the 5% level and all are jointly significant. Consistent with the findings in Table 2, men with better-educated parents have higher earnings, while those from larger families have lower earnings on average. The relationship between sibship size and log earnings appears to depend on the gender composition: holding family size constant, men with more sisters typically earn less, even though the differences are not statistically significant at conventional level. These patterns of parental education and gender composition typically holds true for the within-cohort regressions as well, with smaller estimated effects for the men born between 1940-49. The school quality indicators are always jointly significant, although for some of the cohorts the effects are individually insignificant. Nevertheless, the point estimates indicate that individuals who received better primary and secondary education have higher earnings on average.

[16] Table A.1 in the appendix reports sample averages separately for each cohort.

The evidence contained in Tables 2 and 3 suggests three clear patterns for the effects of family background and school quality on educational attainment and earnings. First, individuals with better-educated parents have higher educational attainment and earnings. Second, men from larger families have less schooling and lower earnings on average. Moreover holding family size constant, those with more sisters have further lower earnings and education. Third, individuals educated in states and cohorts with better school quality (lower pupil-teacher ratio and higher relative teacher salary) also have more schooling and higher earnings, though the relationship is not as strong as it is for family background. Most of these patterns hold true within and across cohorts and are robust to a wide variety of specifications. The next section will exploit and further analyze these patterns in order to identify the effect of family background on the return to education.

3. Theoretical Framework

The objective of this study is to identify the effect of family background on the return to education. With this objective in mind, I set out a formal model of schooling and earnings that explicitly specifies the connection between family background factors, schooling, earnings, and returns to schooling. In light of recent studies of the causal effect of education, the model allows the return to schooling to vary across individuals. Unlike previous studies, however, the model considers the effect of family background on both the marginal benefit <u>and</u> the marginal cost of an additional year of education. Therefore, family background characteristics will affect the return to education in two distinct ways: First, by directly affecting the benefits to an additional year of education, for example through innate ability. Second, families can have some bearing on the amount of schooling acquired by individuals through higher benefits or lower costs, indirectly affecting the marginal return to schooling. A key point is that unless the marginal benefit to an additional year of education is constant, both effects must be considered in order to assess the effect of family background on the return to schooling. Consequently, the parameters of the family background gradients in the marginal benefit and marginal cost schedules must be estimated. To facilitate the discussion of the econometric issues involved and illustrate the implications of the model, suppose that log earnings are determined by the following equation:

$$\log y_i = a_i + b_i S_i - 0.5 k_1 S_i^2 + \gamma' F_i + \delta' Q_i + \varepsilon_i \tag{1}$$

where S_i represents years of completed education, F_i is a measure of family background, and Q_i is a measure of school quality.[17] According to this specification, family background and school quality can directly influence the levels of earnings. In this model, there are two sources of unobserved heterogeneity in log earnings. The intercept a_i represents the level of ability of individuals that does not interact with the level of schooling (i.e. the absolute advantage). The other ability factor b_i, is the heterogeneous component of the education slope interacting with the level of schooling and granting higher net returns to schooling to individuals with higher b_i (i.e. the comparative advantage). Both a_i and b_i represent an unspecified combination of individual specific abilities, influences of familial environment, and inherited skills. In the specification below, both ability factors will be allowed to be freely correlated with family background characteristics. Consistent with the earnings function in (1) are linear marginal benefit and marginal cost schedules with heterogeneous intercepts[18]:

$$\begin{aligned} MB_i &\equiv b_i - k_1 S_i \\ MC_i &\equiv r_i + k_2 S_i \end{aligned}$$

where k_1 and k_2 are positive constants representing the slopes of the schedules, and r_i is a person-specific discount rate. Again, r_i will be allowed to be freely correlated with family background characteristics. Equating marginal benefit and marginal cost yields an expression for optimal

[17] Other determinants of log earnings like labor market experience and race are ignored to keep the presentation simple. In the empirical analysis F_i and Q_i will be vectors of family background and school quality measures.

[18] See Heckman and Vytlacil (1998) and Card (1999)

schooling level:

$$S_i = \frac{b_i - r_i}{k} \qquad (2)$$

where $k = k_1 + k_2$. Equation (2) illustrates that schooling is determined by comparing the marginal benefit and marginal cost of an additional year of schooling. Clearly, in this model, family background affects schooling through its effects on b_i and r_i. To proceed, suppose that the absolute advantage of individuals (a_i), and the individual-specific components of the marginal benefit and marginal cost (b_i and r_i) of an additional year of schooling are related to family background by the following equations:

$$a_i = a_0 + a_1(F_i - \overline{F}) + v_{1i} \qquad (3.1)$$

$$b_i = b_0 + b_1(F_i - \overline{F}) + b_Q(Q_i - \overline{Q}) + v_{2i} \qquad (3.2)$$

$$r_i = r_0 + r_1(F_i - \overline{F}) + r_Z(Z_i - \overline{Z}) + v_{3i} \qquad (3.3)$$

Equations (1)-(3) provide a generalized version of the standard causal model for schooling and log earnings.[19] In the standard model, the unobserved determinants of log earnings and schooling a_i, b_i and r_i are treated solely as random variables, and the parameters a_1, b_1, r_1 and b_Q are all equal to 0. In the present context, b_0 denotes the average causal effect of schooling on log earnings, r_0 is the average discount factor in the population, while b_1 and r_1 measure the effect of family background on the marginal benefit and marginal cost of schooling.

In what follows, the assumption of linearity of the conditional expectations in (3.1)-(3.3) will be maintained: $E[v_{1i}|F_i, Q_i, Z_i] = E[v_{2i}|F_i, Q_i, Z_i] = E[v_{3i}|F_i, Q_i, Z_i] = 0$.[20] This does not rule out that the stochastic components v_{1i}, v_{2i} and v_{3i} are uncorrelated with other variables, in particular with schooling. In this model the conventional ability bias arises because of a correlation between v_{1i} and S_i, while the endogeneity or comparative advantage bias arise because

[19] The model in (1)-(3) is consistent with an optimizing model of schooling choice. See Willis and Rosen (1979), Willis (1986), Card (1999).

[20] An alternative approach would be to make a distributional assumption for the joint distribution of v_{1i}, v_{2i}, v_{3i}. See for example Cameron and Taber (2000), and Taber (2001).

of a correlation between v_{2i} and S_i.

According to the model in (3), Z_i is an instrumental variable for schooling in equation (1): it is an observable variable that affects schooling choices (through (2)), but is uncorrelated with the heterogeneity factors a_i and b_i. Therefore, instrumental variables like can be used to identify the parameters of the marginal benefit schedule. The model embodies another exclusion restriction: conditional on measures of family background, the observable variable Q_i affects the marginal benefit of schooling but not the marginal cost. Therefore, Q_i is not a proper instrumental variable for schooling since it directly influence the return to schooling. However, it is shown below that variables satisfying this kind of exclusion restriction can be used to identify the parameters of the marginal cost schedule.

Substituting equations (3.2) and (3.3) in the schooling equation (2) yields an equation for the realized schooling levels as a function of family background, and the variables excluded from the marginal benefit and marginal cost schedules (Z_i and Q_i):

$$S_i = \frac{1}{k}[(b_0 - r_0) + (b_1 - r_1)(F_i - \overline{F}) + b_Q(Q_i - \overline{Q}) - r_Z(Z_i - \overline{Z}) + v_{2i} - v_{3i}]$$

Therefore the effect of a family background variable F_i on the marginal return to schooling is given by:

$$\begin{aligned} \frac{\partial MB_i}{\partial F_i} &= \frac{\partial b_i}{\partial F_i} - k_1 \frac{\partial S_i}{\partial F_i} \\ &= b_1(1-\omega) + r_1\omega \end{aligned}$$

where $\omega = k_1/k$. This expression shows that the heterogeneity in the return to schooling arise because of the variation in b_i <u>and</u> r_i. Family background can affect the return to schooling through its "direct" effect on b_i (captured by b_1) and its "indirect" effect on amount of schooling acquired, via b_i and r_i (captured by $\omega(b_1 - r_1)$). Therefore the direct and indirect effects (i.e. the parameters b_1, r_1 and ω) must be estimated to fully assess the effect of family background on the return to schooling. In the special case where the earnings function (1) is linear in schooling, the marginal benefits of schooling is constant (which implies $\omega = 0$), and it is sufficient to measure the effects of family background on b_i only. This is the case considered by Altonji and Dunn (1996) and Ashenfelter and Rouse (1998) who obtain mixed evidence on whether parental education affects the return to education. This paper generalizes their studies to the case where family background is allowed to jointly influence education levels and return to education.

4. Empirical Framework

A. *Identification of the family background gradients*

This section shows that when the 2 exclusion restrictions in (3.1)-(3.3) are satisfied, all the parameters of the model in (1)-(3) are identified.[21] The model in (1)-(3) implies the following reduced-form regression for schooling:

$$\begin{aligned} S_i &= E[S_i|F_i, Q_i, Z_i] + \xi_i \qquad (4) \\ S_i &= \pi_{10} + \pi_{11}(F_i - \overline{F}) + \pi_{12}(Q_i - \overline{Q}) + \pi_{13}(Z_i - \overline{Z}) + \xi_i \end{aligned}$$

where:

$$\pi_{10} = \tfrac{b_0 - r_0}{k} \qquad \pi_{12} = \tfrac{b_Q}{k}$$
$$\pi_{11} = \tfrac{b_1 - r_1}{k} \qquad \pi_{13} = \tfrac{-r_Z}{k}$$

Clearly, the parameters measuring the effects of family background on the return to schooling (b_1 and r_1) cannot be estimated from this single regression. The regression function for log earnings is given by:

[21] In the case where one of the exclusion restrictions is not satisfied, the model is under-identified by one parameter, except if the earnings function in (1) is linear. See Deschenes (2001) for an alternative estimation strategy in this case.

$$E[\log y_i|S_i, F_i, Q_i, Z_i] = E[a_i|S_i, F_i, Q_i, Z_i] + E[b_i|S_i, F_i, Q_i, Z_i]S_i - 0.5k_1S_i^2 + \gamma' F_i + \delta' Q_i$$

It follows from (4) and the assumption of linear conditional expectations embodied in equations (3.1)-(3.3) that:

$$E[v_{1i}|S_i, F_i, Q_i, Z_i] = \lambda_S \xi_i \quad (5.1)$$

$$E[v_{2i}|S_i, F_i, Q_i, Z_i] = \Psi_S \xi_i \quad (5.2)$$

where λ_S and Ψ_S are the linear projections coefficients of v_{2i} and v_{3i} on ξ_i.[22] Therefore, the conditional expectations of the unobserved heterogeneity factors a_i and b_i are given by:

$$\begin{aligned} E[a_i|S_i, F_i, Q_i, Z_i] &= a_0 + a_1(F_i - \overline{F}) + \lambda_S \xi_i \\ E[b_i|S_i, F_i, Q_i, Z_i] &= b_0 + b_1(F_i - \overline{F}) + b_Q(Q_i - \overline{Q}) + \Psi_S \xi_i \end{aligned}$$

In this model λ_S is the conventional ability bias due to a correlation between a_i and S_i, and Ψ_S is an endogeneity (or comparative advantage) bias due to a correlation between b_i and S_i.[23] Substituting in the regression function we obtain the following estimating equation:

$$\begin{aligned} \log y_i = {} & \pi_{20} + \pi_{21} S_i + \pi_{22} F_i + \pi_{23} Q_i + \pi_{24} \xi_i + \pi_{25} S_i^2 + \pi_{26} S_i \xi_i + \\ & \pi_{27} S_i F_i + \pi_{28} S_i Q_i + e_i \end{aligned} \quad (6)$$

[22] This derivation follows from the following property of linear least squares projections: $E^*[y|x,z] = E^*[y|x] + \rho[z - E^*[z|x])$. See Sargent (1987).

[23] Note that Ψ_S is related to the fraction of the total variance in schooling outcomes attributable to the heterogeneity in the schooling slopes. Letting f denote this (unobservable) fraction, it can be shown that $\Psi_S = kf$, where k is the sum of the slopes from the marginal benefits and marginal cost equations $(k_1 + k_2)$. Thus, given estimates of k and Ψ_S can be estimated.

where:

$$\pi_{20} = a_0 - a_1\overline{F} \qquad \pi_{25} = -0.5k_1$$
$$\pi_{21} = b_0 - b_1\overline{F} - b_Q\overline{Q} \qquad \pi_{26} = \Psi_S$$
$$\pi_{22} = \gamma + a_1 \qquad \pi_{27} = b_1$$
$$\pi_{23} = \delta \qquad \pi_{28} = b_Q$$
$$\pi_{24} = \lambda_S$$

The model implies a log earnings regression with main effects in years of schooling, family background and school quality, a quadratic term in years of education, interactions of family background and years of education, interactions of years of education and school quality, and finally, a linear term in the schooling reduced-form residual and an interaction between the residuals and schooling.[24] The regression coefficients in (4) and (6) identify all the parameters of the model: the population averages b_0 and r_0, the family background gradients b_1 and r_1, and the slopes of the schedules, k_1 and k_2. The inclusion of ξ_i and $\xi_i S_i$ as controls in the regression will eliminate any ability or endogeneity biases from the relationship between log earnings and years of education.[25] The regression coefficient associated with the interaction between schooling and family background in the log earnings regression identifies b_1, while the regression coefficient associated with the interaction between schooling and the school quality identifies b_Q. The quadratic schooling term provides an estimate of k_1, the slope of the marginal benefit schedule. Given those estimates, the average causal effect of education b_0 is obtained from the main effect coefficient in schooling, while the average discount factor r_0 is obtained from the intercept in the schooling reduced form. Using the estimates from the log earnings regression, it follows from

[24] Dearden, Ferri and Meghir (2002) use a similar wage equation with nonlinear terms in school quality to identify the effect of school quality on wages in a model where the impacts of school quality are heterogeneous.

[25] This "control-function" approach is due to Garen (1984). See Heckman and Vytlacil (1998) and Card (1999) and Woolridge (2000) for discussions of this approach in the context of schooling models.

equation (4) that:

$$
\begin{aligned}
r_0 &= \pi_{21} + \pi_{27}\overline{F} + \pi_{28}\overline{Q} - \frac{\pi_{28}}{\pi_{12}}\pi_{11} \\
r_1 &= \pi_{27} - \frac{\pi_{28}}{\pi_{12}}\pi_{11}
\end{aligned} \tag{7}
$$

B. *Estimation*

This study will focus on the effects of 3 measures of family background on the return to schooling: father's education, mother's education and number of siblings (denoted by F_1, F_2 and F_3). In most data sets, these 3 measures of familial environment are the strongest predictors of children's future outcomes.[26] Two measures of school quality, the pupil-teacher ratio and the relative teacher pay (denoted by Q_1 and Q_2) will be used as exclusion restrictions in the marginal cost equation. Therefore, a maintained assumption in this paper is that holding family background and size constant, measures of school quality at the cohort and state level only influence the marginal benefit of schooling and have no effect on the marginal cost.[27] This exclusion restriction identifies the parameters of the marginal cost equation (i.e. r_0, r_1 and k_2). Following Butcher and Case (1994), measures of siblings gender composition (conditional on family size) will be used as instrumental variables for schooling (i.e. variables that affect schooling only through the cost of schooling). This exclusion restriction identifies the parameters from the marginal benefit equation (i.e. b_0, b_1 and k_1).[28]

The procedure outlined in Section 4a indicates how to interpret the interaction coefficients in the log earnings regression and main effect coefficients in the schooling regression when scalar measures of family background and school quality are included. In the application below, 3

[26] Family income is also strongly associated with children's future outcome, but it is not used in the present analysis because of its lower reliability and potential endogeneity. The results are essentially unchanged when family income at age 16 is also included in the analysis.

[27] This assumption would fail, for example, if an increase in school quality raises the perceived cost of schooling because of the higher effort required to progress through the academic curriculum.

[28] Some tests of the validity of this exclusion restriction are discussed in Section 6.

measures of family background will be considered. I now briefly describes the estimation procedure for this case. Identification requires at least one observable variable that can be excluded from the marginal cost and marginal benefit equations. When more than one variable can be excluded from either equations, the model is over-identified, as is the case here. First a system of 2 regressions is estimated jointly:

$$\begin{aligned} S_i &= \pi_{10} + \pi_{11}(F_{1i} - \overline{F}_1) + \pi_{12}(F_{2i} - \overline{F}_2) + \pi_{13}(F_{3i} - \overline{F}_3) \qquad (8) \\ &\quad + \pi_{14}(Q_{1i} - \overline{Q}_1) + \pi_{15}(Q_{2i} - \overline{Q}_2) + \pi_{16}(Z_i - \overline{Z}) + g(X_{1i}, \gamma_1) + \xi_i \\ \log y_i &= \pi_{20} + \pi_{21} S_i F_{1i} + \pi_{22} S_i F_{2i} + \pi_{23} S_i F_{3i} + \pi_{24} S_i Q_{1i} + \pi_{25} S_i Q_{2i} + \qquad (9) \\ &\quad \pi_{26} S_i + \pi_{27} S_i^2 + \pi_{28} \xi_i + \pi_{29} S_i \xi_i + g(X_{2i}, \gamma_2) + \varepsilon_i \end{aligned}$$

Both X_{1i} and X_{2i} contain 3 cohort dummies, a race indicator, 3 region of birth dummies, 3 indicators for the region of residence in 1973, and an indicator for residence in a metropolitan area (SMSA) in 1973. In addition to these regressors, X_{2i} includes a quartic in labor market experience, and the main effects of the family background and school quality variables. The parameters are obtained by using optimal minimum-distance (OMD) estimation.[29] In the present context, the OMD procedure seeks estimates of the 13 parameters of the marginal cost and marginal benefit schedules that are as close as possible to the predictions of the model, based on the 14 relevant regression coefficients (π_{11}, π_{12}, ..., π_{29}) from equation (8) and (9). The details of the estimation procedure are presented in the appendix.

5. Results

A. *Educational attainment and gender composition among siblings*

Before proceeding to the estimation of the parameters of the model, I further examine the rela-

[29] See Chamberlain (1984).

tionship between educational attainment and gender composition of the siblings. Table 4 presents regressions of completed education on the same specifications as Table 2, adding controls for the presence of any sisters, any brothers, and the fraction of females among siblings. This simple analysis might shed some light on the mechanisms through which family composition affects educational attainment. In each specification, the effects of family size are controlled for by a linear main effect, or by including a series of unrestricted dummies.[30] The entries in Table 4 again provide clear evidence that holding family size and background constant, men who grew up with at least one sister have a significant 0.17-0.20 years of education less. In all specifications, the "any sisters" variable is a stronger predictor of educational attainment than parental education (it is also has a larger effect than number of siblings in the linear specifications). This pattern is contrary to that of Butcher and Case (1994) who found that sibship gender composition had no effect on the educational attainment of men.[31]

As column (3) and (4) indicate, however, conditional on family size and background, there are no differences in the educational attainment of men who grew up with at least one brother and those who did not: the estimated "any brothers" effect is small and insignificant. Moreover, the estimated difference is positive or negative, depending on the specification of the family size effects. Column (5) and (6) show that the fraction of females in the sibship is also an important determinant of the educational attainment of men. In column (7) and (8) both the "any sisters" and the percent female variables are included to determine which one has the strongest effect on educational attainment. It is apparent that the effect of sibling's gender composition on the educational attainment of males is mainly working through the presence of sisters in the sibship. In both specifications, the percent female variable is never significant, and is smaller in magnitude. The indicator for at least one sister has the same magnitude as in

[30] In the OCG samples, the number of siblings variable ranges between 1-19 (mean=3.9), therefore the unrestricted effects specification includes 19 dummies.

[31] The results of Butcher and Case (based on the 1985 wave of the PSID) indicate that conditional on family size, men with at least one sister have more years of education (0.05), but this effect is imprecisely estimated (standard error=0.15). They also note that: "For men in older age cohorts, completed education appears to be negatively related to the presence of any sisters if one does not control adequately for the number of siblings in the family." Even with flexible controls for number of siblings, the OCG data shows a strong negative relationship between educational attainment and the presence of any sisters.

column (1) and is relative precisely estimated (-0.21, with a t-statistic of 2.5) in the linear effects specification, while it is less precisely estimated in the unrestricted effects specification (-0.12, with a t-statistic of 1.4). Overall, the estimates in Table 4 demonstrate that holding family size and background constant, men who grew up with at least one sister have significantly lower educational attainment than those who did not. These patterns are consistent with a model where parents care about the lifetime wealth and labor market earnings of their children, and where the return to educational investments is lower for women (Berhman, Pollak and Taubman 1982). In that case, the presence of sisters in the sibship will be negatively correlated with the educational attainment of males since more family resources will have to be allocated to females in order to equalize labor market earnings. Given that the gender composition of the sibship is random (conditional on other measures of family background), this suggests that indicators of the gender composition can potentially be used as variables affecting schooling only through their effects on the cost of schooling.[32]

B. *Two-stage least squares estimates of the return to education*

As a prelude to the empirical implementation of the model developed in sections 3 and 4, Table 5 presents a series of reduced-form regressions for schooling and log earnings, as well as the OLS and TSLS estimates of the return to education, using the presence any sisters as an instrumental variable for years of completed education. All specifications are based on the same set of controls as Tables 2-4, plus a quartic in labor market experience.[33] The second and third columns of

[32]One drawback of such an instrument is that fertility is a choice variable. Some theoretical models of fertility predict that family size and child ability are negatively related. Since the probability of having at least one sister increases with family size, it is possible that the any sister variable might still be correlated with the unobserved ability factors even after controlling for family size. See Rosenzweig and Wolpin (2000) for a discussion along those lines. Section 6 will present some tests of the validity of the "any sisters" variable as an instrumental variable for years of education.

[33]This measure of labor market experience is based on the reported year of permanent transition to the labor market by the respondent. This measure is used instead of potential labor market experience (age-education-6) to avoid the introduction of additional endogeneity biases. A regression of "actual" on "potential" experience has a slope of 0.88 with an R^2 of 0.83. Note that the models in Table 4 do not include controls for labor market experience. Thus, the inclusion of labor market experience in the models for Table 5 explains the difference in the

Table 5 indicate that holding family size and background constant, men who grew up with at least one sister have lower education (about 0.16 years of education less) and lower earnings (about 3% less).[34] The use of the any sisters indicator as an instrumental variable yields a TSLS estimate of the return to schooling of about 0.20 (with standard error 0.065), which is 3 times as large as the corresponding OLS estimates reported in column 1. This result is consistent with the recent literature. Card (2000) surveys recent studies of the return to education based on instrumental variables. In all studies, the IV estimate is larger than the corresponding OLS estimate, but in most cases the hypothesis that this difference is due to sampling error cannot be rejected. In this sample of men from the OCG, however, the hypothesis that the difference between the OLS and TSLS point estimates is due to sampling error is rejected (p-value=0.03).[35]

C. *Estimates of the family background gradients*

An alternative to TSLS is a control function approach, as in equation (8) and (9). One advantage of the control function approach over TSLS is that it permits the identification of the average causal effect, as well as the identification of the correlation between the treatment variable (schooling) and the unobserved determinants of earnings a_i and b_i that are correlated with the treatment. One drawback from this approach is that it requires stronger assumptions on the nature of the relationship between unobserved ability factors and the observable variables. For the problem at hand, however, a control function approach is more desirable since it allows the direct identification of the family background gradients in the marginal benefit and marginal cost schedules. In addition, the linearity assumptions required for the control function correction to be valid are already embodied in equations (3.1)-(3.3).

Table 6 reports the coefficients from the schooling and log earnings control-function regres-

estimated "any sisters" effect in the schooling reduced-form.

[34] The F-statistic on the excluded instrument in the first-stage is 10.48 (p-value=0.00).

[35] Under the null hypothesis that OLS is consistent, the difference between the OLS and TSLS point estimates divided by the difference in their variances has an asymptotic chi-square distribution. See Hausman (1978) for a presentation of specification tests of this sort.

sions, following the specification of equations (8) and (9). In column (1), coefficients from the reduced-form regression of schooling are reported. Those are essentially the same as the reduced-form coefficients displayed in the first column of Table 5, except that the specification in Table 6 includes main effects in school quality. None of the coefficients are significantly changed by this addition. Column (2) reports the coefficients from the log earnings regression specified in equation (9). In this specification, the residuals from the reduced-form regression of schooling are included in the earnings equation, as are their interactions with schooling. These controls will purge the other regression coefficients of any ability or endogeneity biases. The estimated average return to schooling is 0.16 (with standard error 0.062) is smaller than the TSLS estimate. Note that this estimate is not entirely comparable to the estimate reported in Table 5 since the present specification includes a quadratic in years of education and interactions of family background and education, and school quality and education.

Table 7 presents OMD estimates of the parameters of the marginal benefit (in column 1) and marginal cost schedules (in column 2), derived from the regression coefficients reported in Table 6.[36] For each measure of family background, the "total effect"[37] of that variable on the marginal return to schooling is reported in column (3). The estimated intercepts of each schedule (b_0 and r_0) are displayed in row 1. In row 2-4 are the family background gradients for the marginal benefit and marginal cost equations (3.2) and (3.3). The slopes of each schedule (i.e. k_1 and k_2) are reported in row 5. Finally, rows 6 and 7 show the ability bias term (λ_S) and the comparative advantage selection term (Ψ_S)[38], while row 10 reports the goodness-of-fit statistic associated with the model.

Based on this specification, the estimated average causal effect of education is 0.1618 (with standard error 0.045), which is about 20% smaller than the TSLS estimate. This is not surprising since the framework underlying Table 7 decomposes the causal effect of education into an

[36] Appendix A provides more detail on the OMD estimation of the parameters.

[37] The "total effect" of a family background measure F_j on the marginal return to schooling corresponds to $\partial MB_i/\partial F_{ij} = b_{1j}(1-\omega) + r_{1j}\omega$, for $j = 1, 2, 3$.

[38] Note that this is not the magnitude of the endogeneity bias in the OLS estimate of the returns to schooling. In the standard model, without family background measures entering the b_i equation, the endogeneity bias in the OLS estimate of b correspond to $\Psi_S\overline{S}$.

idiosyncratic component *and* a component due to variation in family background. The average discount rate is 0.0292 (with standard error 0.043). Men with better-educated fathers have higher returns to schooling, ensuing a positive effect of father's education on the heterogeneous component of the marginal benefit, b_i. Conversely, men with better-educated mothers have a lower return to schooling on average, resulting from the large reduction in the marginal cost of schooling associated with mother's education. This finding can also be stated in terms of the characteristics of the production function (1): own education and father's education are q-complements, while own education and mother's education q-substitutes.[39] Finally, men raised in larger families have lower marginal returns to schooling, a result entirely attributable to lower b_i in larger families, conditional on parental education. Interestingly, the results in Table 7 indicate that the measure of family background with the largest effect on the return to schooling is the number of siblings.[40] This follows because the positive effect of father's education on the return is essentially offset by the negative effect mother's education. The estimated slope of the marginal benefit schedule k_1 is essentially 0, which suggest that the marginal return to schooling is roughly constant. As expected the slope of the marginal cost is positive and much steeper.[41] In these data the ability bias is negative and relatively important in magnitude at -0.12 (with standard error 0.043), while the comparative advantage selection term (the projection coefficient of b_i on S_i) is 0.0014.[42] These results are consistent with a model of non-hierarchical sorting: individuals with higher absolute ability levels acquire less schooling, while individuals with higher benefits to schooling acquire more schooling.[43] Finally, the goodness-of-fit statistic

[39] Two factors of productions are said to be q-complements (substitutes) if their partial elasticity of complementarity is positive (negative). See Sato and Koizumi (1973).

[40] An important implication of these results is that measures of family background and parental education are not valid instrumental variables for years of education since they affect the marginal benefit of an additional year of education.

[41] Given that the marginal benefit schedule is almost horizontal, the marginal cost of schooling must be increasing in schooling to ensure interior solutions.

[42] These results are entirely compatible with the OLS estimates of return to schooling reported in Table 5. It can be showed that in this model (under the assumption of no measurement error in schooling), the OLS estimate of the schooling slope in Table 5 converges to $b_0 + \lambda_S + \Psi_S \overline{S}$ = 0.15-0.11+0.0014×12.15 ≈ 0.06.

[43] Willis and Rosen (1979) and Garen (1984) also find evidence of non-hierarchical sorting in the NBER-Th and NLS samples. They focus only on the parameters b_0, λ_S and Ψ_S and their estimation procedure is based on the exclusion restriction that family background variables do not affect the marginal benefit to schooling. Therefore

for this model is 8.52, which slightly higher than the $\chi^2_{(1)}$ critical value, suggesting that this model and its embodied exclusion restrictions provide a too simplistic statistical representation of the data.

D. *Interpretation of the results*

Taken as a whole, the results in this paper provide new evidence on observable and unobservable sources of variation in the return to education. Contrary to the results of Ashenfelter and Rouse (1998) and Altonji and Dunn (1996), the results in this paper indicate that family background variables play an important role in generating variation in the return to schooling across individuals. Allowing the returns to education to vary with family background variables reduces the estimated average causal effect of education by 20%.

Moreover, the results provide clear evidence on the relative importance of the different sources of heterogeneity in explaining schooling outcomes. Based on the results in Table 7, the fraction of the total variance in schooling (9.86) attributable to variation in ability as opposed to variation in the cost (or tastes) for schooling is 11%.[44] In other words, for these data, most of the difference in educational attainment across individuals can be attributed to differences in the cost of an additional year of education.

Finally, there is no evidence that the benefits to an additional year of education are declining with the level of education. The entries in Table 7 suggest that the slope of the marginal benefit schedule is essentially zero. Combined with the negative ability bias reported in Table 7, this suggests a novel interpretation of the recent findings from studies of the causal effect of education based on instrumental variables. Similar to the results in Table 5, these studies (see Card 2000 for a survey) have documented IV estimates of the return to schooling systematically exceeding the corresponding OLS estimates. The leading explanations for this pattern are: (i)

their results are not entirely comparable to those reported in this paper. See Willis (1986) for a discussion of sorting models of education and earnings.

[44] Using the fact that $\Psi_S = kf$, the entries in Table 7 imply that $f = 0.0014/0.0131 = 11\%$.

small ability bias combined with a downward bias in the OLS estimate due to measurement error in reported schooling (Griliches 1979, Angrist and Krueger 1991); (ii) heterogeneity in the returns to education (along with declining marginal benefit to educational investments) combined with instrumental variables that affect the schooling outcomes of individuals who would have relatively low schooling in absence of the supply-side innovation (see e.g. Angrist and Imbens 1995). The results in Table 7 are not consistent with these two explanations. First, attenuation bias alone cannot explain the large gap between the OLS and TSLS estimates in Table 5: Using re-interview data, Bielby, Hauser and Featherman (1977) report that the reliability of reported schooling is about 94% in the OCG.[45]. Second, the marginal returns to education are essentially constant across education levels (i.e. $k_1 \approx 0$). Therefore, as implied by the results of Table 5 and 7, one novel explanation for the larger IV estimates is a negative ability bias in the OLS estimates. Under the assumptions of section 3, it can be shown that the OLS estimate reported in Table 5 converges to $b_0 + \lambda_S + \Psi_S \overline{S}$, i.e. the OLS estimate is confounded by an ability bias (λ_S) and a positive self-selection bias ($\Psi_S \overline{S}$). As long as $|\lambda_S| > |\Psi_S \overline{S}|$, a negative value for λ_S implies that the simple OLS estimate is biased downward.[46] With the relatively high reliability of reported schooling in these data (94%), and essentially no concavity in the "structural" earnings function, this is the only explanation why the TSLS estimates are about 3 times as large as the OLS estimates of the return to education.

6. Robustness of the Results

A. *Validity of the exclusion restrictions*

As discussed in the previous section, siblings gender composition (conditional on family size) can be rightfully excluded from the marginal benefit equation (3.2) if it affects educational attainment

[45] See section 6b for a more detailed discussion of the study by Bielby, Hauser and Featherman.

[46] The results in Willis and Rosen (1979) and Garen (1984) are consistent with a model where $\lambda_S < 0$ and $\Psi_S > 0$.

but has no independent effect on earnings. If gender composition is still related to unobserved determinants of earnings ability after controlling for family background and size, then it does not satisfy the exclusion restriction. While this assumption is not directly testable (since a_i, b_i and r_i) are all unobservable), various pieces of evidence can be examined to evaluate the validity of the exclusion restriction.

Table A.2 in the appendix provides some evidence that conditional on measures of family background, the presence of any sister is unrelated to ability. This table reports the coefficient on an IQ (and other measures of test scores) from a series of regression models fit to the same specifications as in Table 4, using data on cohorts of men from the Project Talent database.[47] The estimated effects of siblings gender composition on the various test scores are small in magnitude and insignificant. Therefore, Table A.2 provides no evidence against the hypothesis that the siblings gender composition (i.e. the any sister variable) is uncorrelated with unobserved ability determinants of earnings (a_i and b_i).

The exclusion restriction on siblings gender composition can also be tested if an additional instrumental variable is available by including the "any sisters" in the earnings equation. An interaction between an index of "poor" family background and the presence of any sisters can be used as an additional instrumental variable. This approach is motivated by the fact that the negative effect of the presence of any sisters on the educational attainment of men (conditional on family size) should be larger for poorer households if it only operates through the marginal cost of schooling.

A continuous index of family background is constructed by the regressing family income when the respondent was 16 on indicators of family characteristics,[48] race, and age, and then using the predicted values from the regression. Individuals with predicted family income smaller than the

[47] Project Talent is a large-scale survey of 5% of all children enrolled in grades 9-12 in 1960, with follow-ups at regular intervals afterwards. At the baseline, demographic and family background information was collected, as well as scores on a battery of aptitude tests. For more information on Project Talent, see "The Project Talent Data Bank: A Handbook," April 1972, American Institutes for Research, Palo Alto, CA.

[48] The family background determinants used in the regression are separate measures of parental education, family size, and indicators of family structure at age 16. The regressions also include 3 cohort dummies, 3 region of birth dummies, and indicators for the region of residence and SMSA status in 1973.

first quartile of the distribution of predicted family income (q=$3162) are classified as having of "poor" family background. Under the assumption that the direct effect of the "any sisters" variable on earnings does not vary by family background, the interaction between this indicator of "poor" family background and the presence of any sisters is a valid instrument for years of education.

Table A.3 presents the reduced-forms regressions and corresponding TSLS coefficients based on this additional instrumental variable. The reduced form coefficients in column (1) confirm that the (negative) effect of the "any sisters" variable on educational attainment are larger for individuals with poorer family background (the interaction term is -0.69, with standard error=0.08). The TSLS estimate of the return to schooling is reported in column (3) along with the estimated direct earnings effect of the any sisters variable. The estimated return is slightly smaller and less precisely estimated than the TSLS estimate reported in Table 7. Table A.3 also confirm that the presence of any sisters has a very small and insignificant effect on log earnings (-0.0029, with a standard error of 0.0107). Again, this provides no evidence against the assumption that holding family size and background constant, the gender composition among siblings is unrelated with the unobserved determinants of earnings.

B. *Measurement error in schooling and parental education*

It is well known that TSLS estimates are not affected by classical measurement error.[49] Less is known, however, on the effects on classical measurement error in nonlinear models, and few analytical formulas describing the bias are available.[50] In the context of the model presented in Section 3, the identification of the parameters of the marginal benefit schedule requires consistent estimates of the interactions terms between family background and schooling in the log earnings regression. Conversely, the identification of the parameters of the marginal cost schedule, based

[49] Kane, Rouse and Staiger (1999) observe that if measurement errors are more closely related to true schooling for the group affected by the instrument, then TSLS may be biased.

[50] See for example Griliches and Ringstad (1970) and Hausman et al. (1991).

on the schooling reduced-form, is not affected by classical measurement error in schooling.[51] This section will study the consistency of the estimates from the log earnings regression in a simple model with classical measurement error in reported schooling and in a single measure of family background, for example father's education. In particular, consider the following true log earnings regression:

$$\log y_i = \beta_0 + \beta_1 S_i + \beta_2 F_i + \beta_3 S_i^2 + \beta_4 S_i F_i + \varepsilon_i$$

Even if schooling and father's education are reported with classical measurement error, the non-linearity introduces non-classical measurement error in the regression. The asymptotic bias in the OLS slope estimates, derived in Appendix E, will in general depend on multiple features of the (unobserved) joint distribution of the measurement error components. Nevertheless, given reliability ratios for reported schooling and father's education, simulations can be used to assess the magnitude of these biases. Under the assumption that the specified values for $\beta_1 - \beta_4$ and the reliability ratios are correct, multiplying the OLS estimates in Tables 4-9 by the corresponding ratios of $(\beta^{true}/\overline{\beta}^{sim})$ will eliminate the measurement error bias.

As part of the 1973 OCG design, a random subsample of about 1,000 respondents was selected for a re-interview survey. Three weeks after the mail return of their OCG questionnaires, individuals were contacted by telephone to obtain a second report of selected items on the OCG questionnaire. Bielby, Hauser and Featherman (1977) report sample correlations and moments from the baseline OCG and the re-interview data. The estimated reliability of own schooling and father's education from those figures are respectively 0.94 and 0.93. Nevertheless, in an effort to be conservative, the simulations will be based on a reliability 0.90 for own schooling and 0.85 for father's education. Table A.4 reports the estimated correction ratios and provides more details on the simulations. This simple analysis suggests that the higher order terms are more sensitive than the main effects. In all specification used, the simulation results suggest that the interaction between schooling and father's education is biased downwards, and that the relevant

[51] Classical measurement error in the dependent variable leads to inefficient estimates, but does not affect the consistency.

entries in Tables 4-8 should be inflated by a factor of 1.25.[52] Therefore the results presented in this study should be interpreted as conservative estimates of the effects of family background on the return to schooling.

C. *Robustness of the OMD estimates*

In certain applications, OMD can be biased downwards (in absolute terms) is there is a correlation between the sampling errors in the vector of moments and sampling error of the elements in the weighting matrix (Altonji and Segal 1996). Table A.5 presents evidence that the estimates reported in Table 7 are not affected by this type of small sample bias. Following Altonji and Segal (1996) who concluded that equally-weighted minimum distance (EWMD) dominates OMD, Table A.5 reports 3 different estimates: EWMD, variance-weighted minimum-distance (VWMD), and direct "one-step" NLLS estimates of the parameters. These alternative estimates of the parameters are essentially identical to those reported in Table 7.

7. Conclusion

This paper develops and implements a simple model of schooling and earnings where the return to schooling varies across individuals, and where family background characteristics play a direct role in generating the heterogeneity. The model illustrates the influence of family background variables in the optimizing behavior of individuals, and thus generalizes the standard causal model of schooling and earnings. It is shown that a correct assessment of the relationship between family background and returns to schooling entails the identification of the effect of family environment on both the marginal benefit and marginal cost of schooling. Only with such information can the impact of family background characteristics on subsequent labor market outcomes be determined.

The empirical analysis, based on a large sample from the 1973 Occupational Change in a

[52] The analysis of Altonji and Dunn (1996) suggests that the within family estimate of the interaction between own schooling and parental education should be inflated by 1.36.

Generation survey documents several patterns concerning the relationship between family background, schooling, earnings and returns to schooling. Parental education raises both the educational attainment and the labor market earnings received by individuals. Men from larger families acquire less schooling and have lower earnings. Moreover, the negative effect of family size is shown to vary with the gender composition of the sibship. Holding family background and sibship size constant, men raised with more sisters have lower educational attainment and earnings. These inferences are robust to a wide variety of specifications.

The patterns are then interpreted in the context of the model. The identification of the parameters of the marginal benefit and cost functions requires two exclusion restrictions. This paper considers measures of elementary and secondary school quality as variables affecting only the marginal benefit of schooling, conditional on family background. The randomness embodied in the gender composition among siblings holding family size and background constant is used as an exogenous source of variation affecting only the marginal cost of schooling. The results provide new evidence on the effects of family background on the returns to schooling. Men raised in larger families have a lower return to schooling, each additional sibling reducing the return to schooling by as much as 5% of the conventional Mincerian estimate. The *combined* effects of parental education are more modest. Men who grew up with better-educated fathers have a higher return to education, while those who grew up with better-educated mothers have a lower return to schooling. This finding suggests that own education and father's education are q-complements in the production of earnings capacity, while own education and mother's education are q-substitutes in the production of earnings capacity. Overall, accounting for family background differences reduces the estimate of the average causal effect of education by 20%. The disparity of these results clearly indicates that different aspects of familial environment have different effects on the marginal benefit and marginal cost of schooling. Family size and father's education entirely operate through the benefits to an additional year of education. The negative impact of maternal education on the return to schooling is solely attributable to lower costs per year of education, and hence improved educational prospects. These insights should be a key component of any appraisal of policies targeted at children from disadvantaged families.

Finally, this paper documents new facts about the sources of variation in the schooling outcomes and in the components of the return to education. A new finding is that almost 90% of the total variance in schooling outcomes is attributable to differences in the costs (or tastes) of schooling, as opposed to differences in ability. There is little evidence of declining marginal benefits to an additional year of education (i.e. the human capital production function is linear in years of education). The results of the optimizing model of schooling choices indicate a negative ability bias and a positive self-selection bias (as found by others, e.g. Willis and Rosen 1979). These last two pieces of evidence support a new interpretation of why the IV estimates exceed the OLS estimates of the returns to education: a negative ability bias combined with constant marginal benefit to schooling are the only explanations consistent with the results presented in this paper.

References

Altonji, Joseph G. and Thomas A.Dunn (1996): "The Effect of Family Characteristics on the Return to Schooling," *Review of Economics and Statistics* 78: 665-671.

Altonji, Joseph G. and Thomas A.Dunn (1996): "Using Siblings to Estimate the Effects of School Quality on Wages," *Review of Economics and Statistics* 78: 665-671.

Altonji, Joseph G. and Lewis M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics* 14: 353-366.

Angrist, Joshua D. and Guido W. Imbens. (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90: 431-442.

Angrist, Joshua D. and Alan B. Krueger. (1991): "Does Compulsory School Attendance Affect Schooling and Earnings," *Quarterly Journal of Economics*, 106: 979-1014.

Ashenfelter, Orley and Cecilia E. Rouse. (1998): "Income, Schooling and Ability: Evidence from a New Sample of Twins," *Quarterly Journal of Economics*, 113: 253-284.

Becker, Gary S. (1967): *Human Capital and the Personal Distribution of Income*, University of Michigan Press, Ann Arbor Michigan.

Berhman, Jere R., Robert A. Pollack, and Paul Taubman (1982): "Parental Preferences and Provision for Progeny," *Journal of Political Economy* 90: 52-73.

Bielby, William T., Robert M. Hauser, and David L. Featherman. (1977): "Response Errors of Black and Nonblack Males in Models of Intergenerational Transmission of Socioeconomic Status," *American Journal of Sociology*, 82: 1242-1288.

Blake, Judith (1989): *Family Size and Achievement*, University of California Press.

Butcher, Kristin F. and Anne Case. (1994): "The Effect of Siblings Sex Composition on Women's Education and Earnings," *Quarterly Journal of Economics*, 109: 531-563.

Cameron Stephen, and Christopher Taber (2000): "Borrowing Constraints and the Returns to Schooling," NBER Working Paper No. 7761.

Card, David and Alan B. Krueger (1992): "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100: 1-40.

Card, David (1993): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," NBER Working Paper No. 4483.

Card, David (1999): "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 14, JAI Press.

Card, David (2000): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," NBER Working Paper No. 7769.

Card, David and Thomas Lemieux (2001): "Dropout and Enrollment Trends in the Postwar Period: What Went Wrong in the 1970s," in Jonathan Gruber, editor,*Risky Behavior Among Youth*, University of Chicago Press.

Chamberlain, Gary (1984): "Panel Data," in Zvi Griliches and Michael Intrilligator, editors, *The Handbook of Econometrics*, volume 2, North-Holland, Amsterdam and New York.

Danziger, Sheldon and Jane Waldfogel (2000): *Securing the Future: Investing in Children from Birth to College*, Sheldon Danziger and Jane Waldfogel editors, Russell Sage Foundation, New York.

Dearden, Lorraine, Javier Ferri and Costas Meghir (2002): "The Effect of School Quality on Educational Attainment and Wages," *Review of Economics and Statistics* 84: 1-20.

Deschenes, Olivier (2001): *An Econometric Analysis of the Returns to Education in the United States*," Ph.D. Thesis, Princeton University, Princeton NJ.

Featherman, David L. and Robert M. Hauser (1978): *Opportunity and Change*, Academic Press.

Garen, John (1984): "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable," *Econometrica* 52: 1199-1218.

Griliches, Zvi and Vidar Ringstad (1970): "Error-in-the-Variables in Nonlinear Contexts," *Econometrica* 38: 368-370.

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

Hauser, Robert M., and Hsiang-Hui Daphne Kuo (1998): "Does the Gender Composition of Sibships Affect Women's Education Attainment?," *Journal of Human Resources* 23: 645-657.

Hausman, Jerry, A. (1978): "Specification Tests in Econometrics," *Econometrica* 46: 1377-1398.

Hausman, Jerry, A., Whitney K. Newey, Hidehiko Ichimura, and James L. Powell (1991): "Identification and Estimation of Polynomial Errors-in-Variables Models," *Journal of Econometrics* 68: 273-295.

Haveman, Robert and Barbara Wolfe (1993): "Children's Prospects and Children's Policy," *Journal of Economic Perspective* 7: 153-174.

Heckman, James J. and Richard Robb (1986): "Alternative Methods for Evaluating the Impact of Interventions," in James J. Heckman and Burton Singer, editors, *Longitudinal Analysis of Labor Market Data*, Cambridge University Press.

Heckman, James J., Anne Layne-Farrar and Petra Todd (1996): "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship," in Gary Burtless editor, *Does Money Matter? The Effects of School Resources on Student Achievement and Adult Success*, Brookings Institution, Washington, DC.

Heckman, James J. and Edward Vytlacil (1998): "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Rate of Return to Schooling When the Return is Correlated with Schooling," *Journal of Human Resources* 23: 974-987.

Kaestner, Robert (1997): "Are Brothers Really Better? Sibling Sex Composition and Educational Attainment Revisited," *Journal of Human Resources* 32: 250-284.

Kane, Thomas, J., Cecilia E. Rouse and Douglas Staiger (1999): "Estimating Returns to Schooling When Schooling is Misreported," NBER Working Paper No. 7235.

Mayer, Suzan E. (1997): *What Money Can't Buy: Family Income and Children's Life Chances*, Harvard University Press.

Mincer, Jacob (1974): *Schooling, Experience and Earnings*, Columbia University Press, New York.

Murphy, Kevin M. and Robert Topel (1985): "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics* 3: 370-379.

Rosenzweig, Mark R. and Kenneth I. Wolpin (2000): "Natural "Natural Experiments" in Economics," *Journal of Economic Literature* 38: 827-874.

Sargent, Thomas, J. (1987): *Macroeconomic Theory*, Academic Press.

Sato, Ryozo and Tetsunori Koizumi (1973): "On the Elasticities of Substitution and Complemetarity," *Oxford Economic Papers* 25: 44-56.

Taber, Christopher (2001): "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability," *Review of Economic Studies* 68: 665-691.

Taubman, Paul (1977): *Kinometrics: The Determinants of Socio-Economic Success Within and Between Families*, Paul Taubman editor, Amsterdam, North-Holland.

Willis, Robert (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Function," in Orley Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics*, North-Holland, Amsterdam and New York.

Willis, Robert and Sherwin Rosen (1979): "Education and Self-Selection," *Journal of Political Economy* 79: S7-S36.

Woolridge, Jeffrey M. (2000): "Instrumental Variables Estimation of the Average Treatment Effect in the Correlated Random Coefficient Model," mimeo, Department of Economics, Michigan State University.

Appendix

A. Optimal Minimum-Distance Estimation

Let π denote the relevant regression coefficients from equations (8) and (9):

$$\pi = [\pi_{10},\ \pi_{11},\ \pi_{12},\ \pi_{13},\ \pi_{14},\ \pi_{15},\ \pi_{21}, \pi_{22},\ \pi_{23},\ \pi_{24},\ \pi_{25},\ \pi_{26},\ \pi_{27},\ \pi_{28},\ \pi_{29}]'$$

When multiple family background and school quality measures are included, the results of section 4a generalizes to:

$$\begin{array}{ll}
\pi_{10} = \frac{b_0 - r_0}{k_1 + k_2} & \pi_{21} = b_{11} \\
\pi_{11} = \frac{b_{11} - r_{11}}{k_1 + k_2} & \pi_{22} = b_{12} \\
\pi_{12} = \frac{b_{12} - r_{12}}{k_1 + k_2} & \pi_{23} = b_{13} \\
\pi_{13} = \frac{b_{13} - r_{13}}{k_1 + k_2} & \pi_{24} = b_{Q1} \\
\pi_{14} = \frac{b_{Q1}}{k_1 + k_2} & \pi_{25} = b_{Q2} \\
\pi_{15} = \frac{b_{Q2}}{k_1 + k_2} & \pi_{26} = b_0 - b_1\overline{F} - b_Q\overline{Q} \\
 & \pi_{27} = -0.5k_1 \\
 & \pi_{28} = \lambda_S \\
 & \pi_{29} = \Psi_S
\end{array}$$

where b_{11} is the element corresponding to F_1 in the vector b_1, r_{11} is the element corresponding to F_1 in the vector r_1, etc. Optimal minimum-distance estimates (OMD) are obtained by minimizing the following quadratic form:

$$\widehat{\theta} = \min[\widehat{\pi} - f(\theta)]'\widehat{W}[\widehat{\pi} - f(\theta)]$$

where $\widehat{\pi}$ is the vector of estimated regression coefficients, $f(\theta)$ is the vector of restrictions imposed by the model on the regression coefficients, $\widehat{W}$ is an estimate of the inverse covariance matrix of $\widehat{\pi}$[53], and θ is the vector of parameters:

$$\theta = [b_0,\ b_{11},\ b_{12},\ b_{13},\ b_{Q1},\ b_{Q2},\ k_1,\ r_0,\ r_{11},\ r_{12},\ r_{13},\ k_2,\ \lambda_S,\ \Psi_S]'$$

Chamberlain (1984) showed that under mild regularity conditions, the optimal minimum-distance estimator is asymptotically efficient. Moreover, the value of the objective function at the optimum can be used to perform specification tests of the model.[54]

[53] Since the system of equation was estimated jointly by GLS, the covariance matrix of π implicitely accounts for the sampling error associated with the estimated residuals, and their interaction with years of education. See Murphy and Topel (1985).

[54] Chamberlain (1984) showed that under the null hypothesis of correct specification, $n[\pi - f(\theta)]'W[\pi - f(\theta)]$ has a χ^2 distribution with $n - k$ degrees of freedom, where $n = \dim(\pi)$ and $k = \dim(\theta)$. Thus, this goodness-of-fit statistic will provides a simple specification test of the model.

B. Measurement error in nonlinear models

This section derives the asymptotic bias in the regression coefficients for a nonlinear model like (5). For simplicity, consider a simple model with classical measurement error in reported schooling and a single measure of family background, for example father's education. In particular, consider:

$$\log y_i = \beta_0 + \beta_1 S_i + \beta_2 F_i + \beta_3 S_i^2 + \beta_4 S_i F_i + \varepsilon_i$$

Suppose that own schooling and father's education are reported with classical measurement error:

$$\begin{aligned} S_i^o &= S_i + v_i \\ F_i^o &= F_i + e_i \end{aligned}$$

where v_i and e_i are independent random variables. Note that this implies non-classical errors in the higher order terms:

$$\begin{aligned} S_i^{o2} &= S_i^2 + 2v_i S_i + v_i^2 \\ S_i^o F_i^o &= S_i F_i + v_i F_i + e_i S_i + e_i v_i \end{aligned}$$

Substituting for the true values, the regression with the observed values is given by:

$$\begin{aligned} \log y_i &= \beta_0 + \beta_1 S_i^o + \beta_2 F_i^o + \beta_3 S_i^{o2} + \beta_4 S_i^o F_i^o + \\ &\quad \{\varepsilon_i - \beta_1 v_i - \beta_2 e_i - \beta_3 [2v_i S_i + v_i^2] - \beta_4 [v_i F_i + e_i S_i + e_i v_i]\} \end{aligned}$$

The asymptotic bias in the OLS estimates is derived by projecting the measurement error components on all the observed variables, S_i^o, F_i^o, S_i^{o2} and $S_i^o F_i^o$:

$$\begin{aligned} v_i &= \rho_{11} S_i^o + \rho_{12} F_i^o + \rho_{13} S_i^{o2} + \rho_{14} S_i^o F_i^o + \eta_{1i} \\ e_i &= \rho_{21} S_i^o + \rho_{22} F_i^o + \rho_{23} S_i^{o2} + \rho_{24} S_i^o F_i^o + \eta_{2i} \\ v_i S_i &= \rho_{31} S_i^o + \rho_{32} F_i^o + \rho_{33} S_i^{o2} + \rho_{34} S_i^o F_i^o + \eta_{3i} \\ v_i^2 &= \rho_{41} S_i^o + \rho_{42} F_i^o + \rho_{43} S_i^{o2} + \rho_{44} S_i^o F_i^o + \eta_{4i} \\ v_i F_i &= \rho_{51} S_i^o + \rho_{52} F_i^o + \rho_{53} S_i^{o2} + \rho_{54} S_i^o F_i^o + \eta_{5i} \\ e_i S_i &= \rho_{61} S_i^o + \rho_{62} F_i^o + \rho_{63} S_i^{o2} + \rho_{64} S_i^o F_i^o + \eta_{6i} \\ e_i v_i &= \rho_{71} S_i^o + \rho_{72} F_i^o + \rho_{73} S_i^{o2} + \rho_{74} S_i^o F_i^o + \eta_{7i} \end{aligned}$$

and substitute back in the regression on the observed values. In general it will the case that:

$$\begin{aligned} \widehat{\beta}_1 - \beta_1 &= f(\beta_1, \beta_2, \beta_3, \beta_4, \rho_{11,}\rho_{12,}...\rho_{17}) \\ &\vdots \\ \widehat{\beta}_4 - \beta_4 &= f(\beta_1, \beta_2, \beta_3, \beta_4, \rho_{41,}\rho_{42,}...\rho_{47}) \end{aligned}$$

Under the assumption that the specified values for $\beta_1 - \beta_6$ and the reliability ratios are correct, multiplying the OLS estimates in Tables 4-7 by the corresponding ratios of ($\beta^{true}/\overline{\beta}^{sim}$) will "correct" the measurement error bias. Table A.3 reports the details of the simulations, as well as the correction ratios.

Table 1: Summary Statistics.

	(I) Full OCG Sample	(II) Full-Time Workers
Age	40.0	41.2
Years of Education	11.75	12.15
Percent Black	0.10	0.10
Weekly Wage	180.9	229.3
Labor Market Experience	20.8	21.3
Number of Siblings	4.01	3.87
Number of Brothers	2.02	1.94
Number of Sisters	1.99	1.93
Father's Education	8.50	8.20
	(0.00)	(0.06)
Mother's Education	9.00	8.72
	(0.00)	(0.08)
Lived on a Farm / Open Country at Age 16	0.19	0.19
Lived with Both Parents at Age 16	0.83	0.83
Born in the South	0.33	0.35
Living in the South in 1973	0.31	0.31
Living in a SMSA in 1973	0.69	0.72
Pupil-Teacher Ratio	---	30.26
Relative Teacher Salary	---	1.07
Observations	32,032	17,300

Note: the entries in parentheses are the fraction imputed.
(I): all observations in the baseline sample of men aged 20-65.
(II): sample of men aged 24-65, earning at least 60$ per week, working at least 48 weeks last year, not in school during the reference week, and with non-missing information on size of the sibship.

Table 2: Reduced-Form Regressions of Educational Attainment on Family Background and School Quality.

	(1)	(2)	(3)	(4)	(5) Born 1910-19	(6) Born 1920-29	(7) Born 1930-39	(8) Born 1940-49	(9) F-Statistics: Equality of (5)-(8)
1. Father's Education	0.1601	0.1458	0.1459	0.1459	0.1870	0.1595	0.1356	0.1328	2.49
	[0.0073]	[0.0073]	[0.0073]	[0.0072]	[0.0183]	[0.0144]	[0.0138]	[0.0125]	(0.06)
2. Mother's Education	0.2274	0.1958	0.1955	0.1948	0.1625	0.1860	0.2252	0.1878	2.53
	[0.0081]	[0.0081]	[0.0081]	[0.0080]	[0.0194]	[0.0150]	[0.0150]	[0.0146]	(0.06)
3. Number of Siblings	---	-0.1662	---	---	---	---	---	---	---
		[0.0071]							
4. Number of Brothers	---	---	-0.1467	-0.1435	-0.2037	-0.1313	-0.1683	-0.0954	3.63
			[0.0121]	[0.0121]	[0.0271]	[0.0237]	[0.0237]	[0.0224]	(0.01)
5. Number of Sisters	---	---	-0.1862	-0.1845	-0.1242	-0.2081	-0.1789	-0.2144	2.49
			[0.0123]	[0.0123]	[0.0275]	[0.0235]	[0.0240]	[0.0236]	(0.06)
6. Pupil-Teacher Ratio	---	---	---	-0.0608	-0.0485	-0.0370	-0.0330	-0.0342	0.38
				[0.0061]	[0.0108]	[0.0105]	[0.0132]	[0.0159]	(0.75)
7. Relative Teacher Salary	---	---	---	0.5922	0.6766	0.3672	0.3500	0.2187	1.69
				[0.0910]	[0.1501]	[0.1190]	[0.1582]	[0.2106]	(0.16)
F-Statistics									
8. Family Background Main Effects	1019.01	950.62	792.97	786.82	165.29	254.59	269.29	205.76	---
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
9. School Quality Main Effects	---	---	---	71.38	24.38	12.02	5.85	3.29	---
				(0.00)	(0.00)	(0.00)	(0.00)	(0.04)	
10. R-squared	0.30	0.32	0.33	0.33	0.31	0.31	0.32	0.28	---

Sample size is 17,300. Standard errors in brackets, p-values in parentheses.
All models include a race indicator, 3 cohort dummies, 3 region of birth dummies, 3 indicators for region of residence in 1973, SMSA status in 1973, and imputation dummies.
Other family background controls are indicators for living with both parents at age 16 and for living on a farm at age 16.

Table 3: Reduced-Form Regressions of Log Earnings on Family Background and School Quality.

	(1)	(2)	(3)	(4)	(5) Born 1910-19	(6) Born 1920-29	(7) Born 1930-39	(8) Born 1940-49	(9) F-Statistics: Equality of (5)-(8)
1. Father's Education	0.0103	0.0094	0.0094	0.0094	0.0099	0.0132	0.0124	0.0048	3.28
	[0.0012]	[0.0012]	[0.0012]	[0.0012]	[0.0029]	[0.0023]	[0.0022]	[0.0020]	(0.02)
2. Mother's Education	0.0163	0.0142	0.0142	0.0146	0.0152	0.0155	0.0167	0.0087	2.36
	[0.0013]	[0.0013]	[0.0013]	[0.0013]	[0.0031]	[0.0024]	[0.0024]	[0.0023]	(0.07)
3. Number of Siblings	---	-0.0108	----	---	---	---	---	---	---
		[0.0011]							
4. Number of Brothers	---	---	-0.0096	-0.0090	-0.0088	-0.0116	-0.0145	-0.0027	1.86
			[0.0019]	[0.0019]	[0.0044]	[0.0038]	[0.0037]	[0.0036]	(0.13)
5. Number of Sisters	---	---	-0.0121	-0.0116	-0.0119	-0.0155	-0.0091	-0.0097	0.59
			[0.0020]	[0.0020]	[0.0044]	[0.0038]	[0.0038]	[0.0038]	(0.62)
6. Pupil-Teacher Ratio	---	---	---	-0.0031	-0.0050	-0.0026	-0.0009	0.0065	4.99
				[0.0010]	[0.0017]	[0.0017]	[0.0021]	[0.0026]	(0.00)
7. Relative Teacher Salary	---	---	---	0.1380	0.1200	0.0850	0.1664	0.1914	4.43
				[0.0146]	[0.0241]	[0.0191]	[0.0254]	[0.0338]	(0.00)
F-Statistics									
8. Family Background Main Effects	197.32	176.89	147.50	145.58	30.70	65.00	62.72	15.44	---
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
9. School Quality Main Effects	---	---	---	49.82	19.61	11.80	21.77	17.45	---
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
10. R-squared	0.18	0.18	0.18	0.19	0.16	0.18	0.22	0.11	---

Sample size is 17,300. Standard errors in brackets, p-values in parentheses.
All models include a race indicator, 3 cohort dummies, 3 region of birth dummies, 3 indicators for region of residence in 1973, SMSA status in 1973, and imputation dummies.
Other family background controls are indicators for living with both parents at age 16 and for living on a farm at age 16.

Table 4: Gender Composition, Family Size, and Educational Attainment.

	(1) Linear Effects	(2) Unrestricted Effects	(3) Linear Effects	(4) Unrestricted Effects	(5) Linear Effects	(6) Unrestricted Effects	(7) Linear Effects	(8) Unrestricted Effects
Any Sisters	-0.2085 [0.0533]	-0.1698 [0.0616]	---	---	---	---	-0.2056 [0.0829]	-0.1224 [0.0902]
Any Brothers	---	---	-0.0668 [0.0535]	0.0488 [0.0630]	---	---	---	---
Percent Female	---	---	---	---	-0.2775 [0.0918]	-0.2437 [0.0981]	-0.0067 [0.1426]	-0.0999 [0.1435]
Father's Education	0.1454 [0.0072]	0.1429 [0.0073]	0.1457 [0.0073]	0.1430 [0.0073]	0.1458 [0.0073]	0.1429 [0.0073]	0.1452 [0.0073]	0.1429 [0.0072]
Mother's Education	0.1955 [0.0081]	0.1947 [0.0080]	0.1958 [0.0081]	0.1948 [0.0081]	0.1956 [0.0081]	0.1946 [0.0080]	0.1956 [0.0081]	0.1947 [0.0080]
Number of Siblings	-0.1532 [0.0078]	Yes	-0.1617 [0.0079]	Yes	-0.1648 [0.0071]	Yes	-0.1468 [0.0084]	Yes
R-Squared	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33

Dependent variable: years of education.
Sample size is 17,300. Standard errors in brackets.
Number of siblings is controlled for by linear main effects (columns 1,3,5,7) or by unrestricted effects (19 dummies, columns 2,4,6,8).
All models include a race indicator, 3 cohort dummies, 3 region of birth dummies, 3 indicators for region of residence in 1973, SMSA status in 1973, and imputation dummies.
Other family background controls are indicators for living with both parents at age 16 and for living on a farm at age 16.

Table 5: Reduced-Form, OLS, and Two-Stage Least Square Estimates of the Return to Schooling.

	(1) OLS	(2) Reduced-Form: Years of Education	(3) Reduced-Form: Log Earnings	(4) TSLS
Any Sisters	-0.0227 [0.0080]	-0.1594 [0.0492]	-0.0313 [0.0084]	---
Years of Education	0.0539 [0.0012]	---	---	0.1963 [0.0667]
Father's Education	0.0029 [0.0011]	0.1116 [0.0067]	0.0090 [0.0011]	-0.0129 [0.0076]
Mother's Education	0.0047 [0.0012]	0.1537 [0.0075]	0.0130 [0.0013]	-0.0172 [0.0104]
Number of Siblings	-0.0007 [0.0012]	-0.1316 [0.0072]	-0.0078 [0.0012]	0.0180 [0.0096]
R-squared	0.28	0.43	0.21	0.14
F-statistic on the excluded instrument	---	10.48 (0.00)	---	---

Sample size is 17,300. Standard errors in brackets, p-values in parentheses.
All models include a race indicator, 3 cohort dummies, 3 region of birth dummies, 3 indicators for region of residence in 1973, SMSA status in 1973, and imputation dummies. Other family background controls are indicators for living with both parents at age 16 and living on a farm at age 16.

Table 6: Estimates of the Schooling Reduced-Form and "Augmented" Log Earnings Regression.

	Years of Education	Log Earnings
Any Sisters	-0.1613 [0.0491]	---
Years of Education	---	0.1567 [0.0619]
Years of Education Squared	---	-0.0001 [0.0008]
Reduced-Form Residuals	---	-0.1188 [0.0597]
Reduced-Form Residuals × Education	---	0.0014 [0.0009]
Father's Education[a]	0.1119 [0.0067]	0.0011 [0.0004]
Mother's Education[a]	0.1548 [0.0075]	-0.0005 [0.0004]
Number of Siblings[a]	-0.1294 [0.0072]	-0.0010 [0.0004]
Pupil-Teacher Ratio[a]	-0.0298 [0.0056]	-0.0003 [0.0002]
Teacher Relative Salary[a]	0.5753 [0.0841]	0.0094 [0.0033]

Sample size is 17,300. Standard errors in brackets, p-values in parentheses.
The standard errors are corrected for the first-stage estimation of the schooling residuals.
See notes to Table 5 for a list of variables included in the regressions.
(a): interacted with education in the log earnings model
The regressions correspond to equations (8) and (9) in the text.

Table 7: OMD Estimates of the Effects of Family Background on the Return to Schooling.

	(1) Marginal Benefit Schedule	(2) Marginal Cost Schedule	(3) Total Effect
1. Intercept	0.1618 [0.0449]	0.0292 [0.0428]	---
2. Father's Education	0.0011 [0.0003]	-0.0004 [0.0003]	0.0011 [0.0002]
3. Mother's Education	-0.0005 [0.0003]	-0.0025 [0.0003]	-0.0006 [0.0003]
4. Number of Siblings	-0.0011 [0.0003]	0.0007 [0.0003]	-0.0010 [0.0003]
5. Slopes (k_1 and k_2)	0.0003 [0.0010]	0.0128 [0.0013]	---
6. Ability Bias (λ_S)	-0.1187 [0.0429]	---	---
7. Self-Selection Bias Term (ψ_S)	0.0014 [0.0005]		
8. Pupil-Teacher Ratio	-0.0004 [0.0001]		
9. Relative Teacher Salary	0.0076 [0.0012]	---	---
10. N×Objective [d.f.] (p-value)	8.52 [1] (0.00)	---	---

Standard errors in brackets, p-values in parentheses.
Based on the estimated coefficients and covariance matrix from Table 6.
See Appendix A for details on the OMD procedure.

Table A.1: Summary Statistics, Sample of Full-Time Workers—by Birth Cohort.

	Born 1910-1919	Born 1920-1929	Born 1930-39	Born 1940-49
Age	58.0	48.4	38.6	28.4
Years of Education	11.03	11.78	12.37	12.87
Percent Black	0.08	0.09	0.10	0.10
Weekly Wage	227.0	249.1	244.6	201.9
Labor Market Experience	38.3	28.3	18.5	8.9
Number of Siblings	4.46	4.08	3.84	3.42
Number of Brothers	2.21	2.03	1.92	1.75
Number of Sisters	2.25	2.04	1.93	1.67
Father's Education	6.85	7.37	8.16	9.61
	(0.11)	(0.09)	(0.09)	(0.08)
Mother's Education	7.11	7.91	8.84	10.13
	(0.10)	(0.07)	(0.07)	(0.05)
Lived on a Farm / Open Country at Age 16	0.23	0.22	0.19	0.14
Lived with Both Parents at Age 16	0.83	0.81	0.84	0.85
Born in the South	0.32	0.35	0.37	0.36
Living in the South in 1973	0.28	0.30	0.32	0.33
Living in a SMSA in 1973	0.72	0.72	0.71	0.71
Pupil-Teacher Ratio	34.24	31.95	29.30	27.55
Relative Teacher Salary	1.09	1.08	1.07	1.07
Observations	2,963	4,553	4,492	5,292

Note: the entries in parentheses are the fraction imputed.
Sample of men aged 24-65, earning at least 60$ per week, working at least 48 weeks last year, not in school during the reference week, and non-missing information on size of the sibship.

Table A.2: Gender Composition, Family Size, and Observable Measures of Ability.

	Dependent Variable = IQ	Dependent Variable = Math Test Score	Dependent Variable = Verbal Test Score
Any Sisters	-0.305	0.345	-0.240
	[0.89]	[0.52]	[0.36]
Father's Education	1.226	1.044	0.544
	[0.11]	[0.06]	[0.04]
Mother's Education	0.986	0.726	0.313
	[0.12]	[0.07]	[0.05]
Number of Siblings	Yes	Yes	Yes
Mean of the Dependent Variable	100.0	100.0	100.0
Observations	28,070	28,070	28,070
R-Squared	0.23	0.19	0.25

Data are from the Project Talent database. Standard errors in brackets.
Number of siblings is controlled for by unrestricted effects (19 dummies).
All models include 3 cohort dummies, indicators for region of residence at the baseline, SMSA status at the baseline.
Other family background controls are indicators for living with both parents and for living on a farm at the baseline.

Table A.3: Reduced-Form, OLS, and Two-Stage Least Square Estimates of the Return to Schooling, Based on the Interaction Between the "Any Sisters" Indicator and Predicted Family Income.

	(1) Reduced-Form: Years of Education	(2) Reduced-Form: Log Earnings	(3) TSLS
Any Sisters	-0.1439 [0.0492]	-0.0285 [0.0084]	-0.0029 [0.0107]
Any Sisters × Low Predicted Family Income	-0.6905 [0.0781]	-0.1230 [0.0134]	---
Years of Education	---	---	0.1781 [0.0232]
Father's Education	0.1042 [0.0068]	0.0076 [0.0012]	-0.0109 [0.0029]
Mother's Education	0.1425 [0.0076]	0.0110 [0.0013]	-0.0144 [0.0039]
Number of Siblings	-0.1266 [0.0072]	-0.0069 [0.0012]	0.0157 [0.0034]
R-squared	0.43	0.21	0.16
F-statistic on the excluded instrument	78.3 (0.00)	---	---

Sample size is 17,300. Standard errors in brackets, p-values in parentheses.
The additional instrument is an interaction between an indicator of "low" predicted family income at age 16 (an indicator for a predicted family income in the lowest quartile), and an indicator for the presence of "any sisters".
All models include a race indicator, 3 cohort dummies, 3 region of birth dummies, 3 indicators for region of residence in 1973, SMSA status in 1973, and imputation dummies. Other family background controls are indicators for living with both parents at age 16 and for living on a farm at age 16.

Table A.4: Monte Carlo Analysis of Measurement Error in Nonlinear Regressions.

	Simulation: Reliability ratio of reported schooling = 0.90; Reliability ratio of reported father's education = 0.85		
	True Values	OLS	True Values/OLS
Design I			
[γ_1=0.08, γ_2=0.0030, γ_3=0.02, γ_4=0.0020]			
Intercept	5.0000	5.1309	0.97
Linear Schooling	0.0800	0.0808	0.99
Linear Father's Education	0.0200	0.0190	1.05
Quadratic Schooling	0.0030	0.0024	1.25
Schooling × Father's Education	0.0020	0.0015	1.33
Design II			
[γ_1=0.08, γ_2=0.0030, γ_3=-0.02, γ_4=0.0020]			
Intercept	5.0000	5.0814	0.98
Linear Schooling	0.0800	0.0806	0.99
Linear Father's Education	-0.0200	-0.0151	1.32
Quadratic Schooling	0.0030	0.0024	1.25
Schooling × Father's Education	0.0020	0.0015	1.33
Design III			
[γ_1=0.05, γ_2=0.0015, γ_3=0.01, γ_4=0.0010]			
Intercept	5.0000	5.0775	0.98
Linear Schooling	0.0500	0.0494	1.01
Linear Father's Education	0.0100	0.0095	1.05
Quadratic Schooling	0.0015	0.0012	1.25
Schooling × Father's Education	0.0010	0.0008	1.25
Design IV			
[γ_1=0.05, γ_2=0.0015, γ_3=-0.01, γ_4=0.0010]			
Intercept	5.0000	5.0515	0.98
Linear Schooling	0.0500	0.0494	1.01
Linear Father's Education	-0.0100	-0.0074	1.05
Quadratic Schooling	0.0015	0.0012	1.25
Schooling × Father's Education	0.0010	0.0008	1.25

The model generating log earnings is:

$\log y = 5.0 + \gamma_1 S + \gamma_2 F + \gamma_3 S^2 + \gamma_4 S \times F + e$
$e \sim N(0\ ,\ 0.2)$
$S \sim N(12\ ,\ 9)$
$S^o = S + v,\quad v \sim N(0\ ,\ \sigma_v^2)$
$F \sim N(8.5\ ,\ 12.25)$
$F^o = F + u,\quad u \sim N(0\ ,\ \sigma_u^2)$

10,000 replications based on samples of size 5,000 were used.

Table A.5: Alternative Estimates of the Effects of Family Background on the Return to Schooling.

	One-Step NLLS		VWMD		EWMD	
	(1) Marginal Benefit Schedule	(2) Marginal Cost Schedule	(1) Marginal Benefit Schedule	(2) Marginal Cost Schedule	(1) Marginal Benefit Schedule	(2) Marginal Cost Schedule
1. Intercept	0.1589 [0.0616]	0.0299 [0.0698]	0.1506 [0.0596]	0.0222 [0.0120]	0.1392 [0.0596]	0.0327 [0.0243]
2. Father's Education	0.0011 [0.0004]	-0.0003 [0.0006]	0.0011 [0.0004]	-0.0003 [0.0004]	0.0006 [0.0004]	-0.0006 [0.0004]
3. Mother's Education	-0.0005 [0.0004]	-0.0025 [0.0007]	-0.0005 [0.0004]	-0.0025 [0.0004]	-0.0010 [0.0005]	-0.0027 [0.0004]
4. Number of Siblings	-0.0010 [0.0004]	0.0006 [0.0006]	-0.0010 [0.0004]	0.0006 [0.0004]	-0.0013 [0.0004]	0.000 [0.0004]
5. Slope	0.0003 [0.0017]	0.0125 [0.0038]	0.0003 [0.0017]	0.0124 [0.0012]	0.0046 [0.0017]	0.0089 [0.0045]
6. Ability Bias (λ_S)	-0.1178 [0.0592]	---	-0.1187 [0.0569]	---	-0.1112 [0.0589]	---
7. Comparative Advantage Selection Term (ψ_S)	0.0014 [0.0009]	---	0.0014 [0.0009]	---	0.0014 [0.0009]	--

Standard errors in brackets, p-values in parentheses.
VWMD and EWMD are based on the estimated coefficients and covariance matrix from Table 6.
For VWMD the weighting matrix is a diagonal matrix with inverse variances as elements on the diagonal. For EWMD the weighting matrix is the identity matrix.
The "one-step" NLLS estimates are obtained by imposing the structure of the model on equations (8) and (9) and fitting the regressions by non-linear least squares.