OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 74, 5 (2012) 0305-9049
doi: 10.1111/j.1468-0084.2011.00667.x

Does Increasing Parents' Schooling Raise the Schooling of the Next Generation? Evidence Based on Conditional Second Moments*

Lídia Farré†, Roger Klein‡ and Francis Vella§

†*Institut d'Anàlisi Econòmica, Campus UAB 08193, Bellaterra, Spain (e-mail: lidia.farre@iae.csci.es)*
‡*Rutgers. The State University, Department of Economics, 75 Hamilton Street, New Jersey, USA (e-mail: klein@economics.rutgers.edu)*
§*Georgetown University, Department of Economics, 37th and O Streets, NW, Washington DC, 20057 USA (e-mail: fgv@georgetown.edu)*

Abstract

This article investigates the degree of intergenerational transmission of education for individuals from the National Longitudinal Survey of Youth 1979. Rather than identifying the causal effect of parental education via instrumental variables we exploit the feature of the transmission mechanism responsible for its endogeneity. More explicitly, we assume the intergenerational transfer of unobserved ability is invariant to the economic environment. This, combined with the heteroskedasticity resulting from the interaction of unobserved ability with socioeconomic factors, identifies the causal effect. We conclude that the observed intergenerational educational correlation reflects both a causal parental educational effect and a transfer of unobserved ability.

I. Introduction

Although it is well established that a positive correlation exists between an individual's educational attainment and that of his/her parents, it remains unknown what it precisely captures.[1] Although some interpret it as a causal relationship, others argue it reflects the intergenerational transfer of unobservable traits. As isolating the causal component of educational transmission is crucial for developing educational related policies it has become an objective of empirical work to appropriately estimate it.

*We are grateful to the editor and two anonymous referees for helpful comments. Farré is a research fellow at IZA, MOVE and Barcelona GSE and acknowledges the support of IVIE (Instituto Valenciano de Investigaciones Económicas), the Barcelona GSE Research Network and the Government of Catalonia, as well as the Spanish Ministry of Science (grant ECO2008-05721/ECON).
JEL Classification numbers: C31, J62.

[1]Haveman and Wolfe (1995) and Behrman (1997) provide extensive surveys of the earlier literature on the intergenerational transmission of education.

To identify this causal component some studies have focused on twins, assuming they have similar values of unobservable traits, and examined the within-twin variation in their educational levels and that of their children. Behrman and Rosenzweig (2002) examine a sample of twins in the US and find a positive effect from the father's education but a small, and possibly negative, effect from that of the mother. However, Antonovics and Goldberger (2005) find this result is sensitive to coding and sample selection rules and conclude that mother's education and father's education do not play dramatically different roles. Studies which use data for adoptees, under the presumption that the 'inheritable traits' are not relevant due to the absence of a genetic relationship between child and parent, find weak effects for the adoptive mother's schooling and large effects for the adoptive father's schooling (Plug, 2004). Björklund, Lindahl and Plug (2006) use information for both the adoptive and biological parents and find that both pre and post birth factors contribute to adopted children's education levels. However, after accounting for assortative mating, through the simultaneous inclusion of both parent's schooling, the effect from the adoptive mother's education vanishes. They find, however, that the education of both adoptive parents is relevant to whether the child obtains university education. This last result is consistent with the evidence in Sacerdote (2004). Black, Devereux and Salvanes (2005) and Chevalier (2004) identify the causal effect by using schooling reforms that produce exogenous variation in the educational choices of parents. These studies find a large positive effect of the mother's education but no significant effect from the father's. This range of conclusions partially reflects the use of different data sets but also highlights that alternative approaches may not identify the causal effect from the same part of the educational distribution. Holmlund, Lindahl and Plug (2011) employ a Swedish data set that allows multiple identification strategies and conclude that the estimated effect of parental education depends on the identification condition employed.

Whereas these existing studies provide important insight each has some limitation. The results for adoptees and twins are based on samples drawn from atypical populations while those which exploit educational reforms identify the causal effect for individuals whose behaviour responds to the reform. We contribute to this debate by providing estimates based on an alternative identifying strategy applied to a more representative sample. We exploit the nature of the intergenerational transmission of unobservable traits to derive a restriction that identifies the causal effect of parental education. Namely, we assume that the correlation of unobservables across generations is invariant to the individuals' socio-economic environments. This assumption seems reasonable when the unobservables are interpreted as inherited ability. In the following section, we describe the model and discuss our identification and estimation strategies. Section III presents the data and our empirical results. Section IV provides some concluding comments.

II. Empirical model

Consider the following model of educational transfer:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + u_i, \quad i = 1, \ldots, N \tag{1}$$

$$S_i^j = X_i\delta_j + v_i^j, \quad j = M, F \tag{2}$$

where S_i^C denotes the child's years of education; S_i^j denotes the parent's years of education (i.e. M for mother and F for father); X_i denotes a vector of exogenous variables which we assume, for generality, to be the same for children and parents; the $\beta's$ and $\delta's$ represent unknown parameters; and the u_i and v_i are error terms with a non-zero covariance which reflects the endogeneity of S_i^M and S_i^F. This non-zero covariance renders the OLS estimates of β inconsistent. As we allow the same X to enter equations (1) and (2) there is no exogenous source of variation in parental education which identifies β. That is, there are no available instruments.[2]

To consistently estimate β we begin by characterizing the structure of the error terms in equations (1) and (2). We first assume that the X_i vector is exogenous. This implies:

$$E[u_i \mid X_i] = E[v_i^j \mid X_i] = 0. \tag{3}$$

The second assumption is that the errors are heteroskedastic. That is, let $H_u^2(X_i)$ and $H_v^{j2}(X_i)$ denote the conditional variance functions for u_i and v_i where:

$$u_i = H_u(X_i)u_i^* \text{ and } v_i^j = H_v^j(X_i)v_i^{*j}, \tag{4}$$

where u_i^* and v_i^{*j} are correlated homoskedastic error terms which we interpret as measures of unobserved ability. Equation (4) indicates that individuals receive a draw from the homoskedastic error distributions denoted by u_i^* and v_i^{*j}. However, because of the presence of heteroskedasticity in the model the contribution of these disturbances to the observed level of education depends on their respective socioeconomic environments or observed characteristics in conjunction with the relevant H function. More precisely, the contribution of the homoskedastic disturbance is scaled up (or down) depending on the form of heteroskedasticity in the model. For example, consider the educational attainment level of the child and assume that u_i^* captures unobserved ability. Suppose that one of the elements of X_i is school type. If the objective of one type of school was to maximize the educational attainment of the better students one could imagine that the school would attempt to identify high ability students and then focus its attention on those students. Alternatively, a different type of school may choose to ensure that lower ability students acquire some minimum level of education. In this setting, it is clear that while u_i^* captures the level of unobserved ability the contribution of this unobserved ability to schooling attainment depends on school type. Moreover, even if school type is included in the conditional mean of educational attainment the resulting disturbances will be heteroskedastic because of this interaction between school type and unobserved ability. Equation (4) captures the manner in which we assume the homoskedastic error terms are scaled.

An implication of equation (4) is that the intergenerational transmission of unobserved ability operates through the relationship between u_i^* and the $v_i^{*j\prime}s$ and not, necessarily, that between u_i and the $v_i^{j\prime}s$. The former captures the manner parents' unobserved ability is transferred to their children while the latter captures how children's and parent's unobserved ability are correlated after each is scaled up by the appropriate H function.

Ordinary least square (OLS) estimation of equation (1) produces inconsistent estimates because of the lack of orthogonality between the $S_i^{j\prime}s$ and u_i and the moments corresponding

[2]Our specification of the conditional mean of the education equations in the empirical section of the article is guided by previous empirical work. Whereas some of the included explanatory variables are possibly endogenous, we do not address this issue and focus our attention solely on the endogeneity of the parent's education variables.

to equation (3) are insufficient to identify the model. Accordingly, we impose two additional conditions which follow from our interpretation of the intergenerational transfer of ability. We impose that the transfer of unobserved ability is independent of the parents' and child's environment. This implies that the conditional partial correlations between the homoskedastic error terms are constant.[3] That is:

$$E[u_i^* v_i^{*j} \mid X_i] = E[u_i^* v_i^{*j}] = \rho^j, \quad j = M, F. \tag{5}$$

Following Klein and Vella (2010) these 'constant conditional partial correlation coefficient' moments can in conjunction with equation (3), and in the presence of equation (4), identify the model.[4]

Using these moments one can estimate the model by generalized method of moments (GMM). However, the estimation of these conditional moments is complicated because of their dependence on the unknown conditional variances and covariances. Klein and Vella (2010) show that the same moments can be imposed by estimating the following control function model:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + \rho^M \frac{H_{ui}}{\hat{H}_{vi}^M} \hat{v}_i^M + \rho^F \frac{H_{ui}}{\hat{H}_{vi}^F} \hat{v}_i^F + e_i, \tag{6}$$

where $\hat{v}_i^M$ and $\hat{v}_i^F$ are the residuals from the parent's education equations; H_{ui} denotes the unknown $H_u(X_i)$ while $\hat{H}_{vi}^j$ are the estimates of $H_v^j(X_i)$; and e_i is a zero mean disturbance which is uncorrelated with the included regressors.[5]

Estimation of equation (6) is considerably simpler than the corresponding GMM procedure but is infeasible here because of the large dimension of X and the unknown nature of the H functions. Klein and Vella (2010) identify the parameters in equation (6) assuming that the $X's$ enter the H functions in an index form but without imposing any structure on the $H's$. Thus their identification results are based on non-parametric and semiparametric representations of the heteroskedasticity. Whereas this is theoretically attractive, as identification is not reliant on specific forms of heteroskedasticity, it is computationally demanding. More importantly, the non and semiparametric approaches have significant data demands as they rely upon the presence of several continuous exogenous variables in the conditioning sets to show identification. To overcome these data demands and to reduce computation the H functions can be parameterized. Accordingly, we specify the following form:

$$H_{ui}^2 = \exp(Z_{ui}\theta_1) \tag{7}$$

$$H_{vi}^{j2} = \exp(Z_{vi}^j\theta_{2j}), \quad j = M, F, \tag{8}$$

[3] Klein and Vella (2010) show that this constant conditional correlation assumption is consistent with a number of data generating processes. They also show that the disturbances may contain more than one component. While this does not invalidate the estimation procedure it may, in some instances, change the 'economic' interpretation of the correlation coefficient.

[4] Klein and Vella (2010) also assume that the ratios (H_{ui}/H_{vi}^M) and (H_{ui}/H_{vi}^F) are not constant across i. This appears to be very mild requirement.

[5] Klein and Vella (2010) only explicitly examine the case of one endogenous regressor. However, as the endogenous regressors are continuous their approach is applicable to the two endogenous regressor case.

where the $Z's$ are the vector of variables considered to be responsible for the heteroskedasticity in the respective equations and the $\theta's$ are unknown parameters to be estimated.[6] The appendix provides a detailed discussion of how the estimator is implemented.[7]

Before proceeding, consider whether the key identification assumptions of this strategy seem reasonable in this context. The first is the presence of heteroskedasticity and there are many reasons why it might occur. If 'distance to school' is a determinant of the level of educational attainment it is likely that an unequal geographical allocation of the number, and quality, of educational institutions may produce important differences in both the mean and variance of educational attainment across regions. Heteroskedasticity may also arise from the heterogenous impact of many of the determinants of education. For example, the cultural diversity of immigrants to the US suggests there are likely to be large differences in the educational attainment of this group. Therefore even after the inclusion of an indicator function capturing that individuals were born overseas, the dispersion in their schooling levels is likely to be different than that for natives.

The second requirement is the constancy of the conditional correlation coefficients. This means that the 'transfer of unobserved ability', measured by the correlation coefficients between u_i^* and the $v_i^{*j\prime}s$, is independent of the socioeconomic environment. This would be satisfied if the transfer reflected some 'genetic' transmission of innate intelligence or ability in the same manner that other genetic endowments, such as skin and eye colour, are transferred from parents to children independently of the economic environment. The assumption would be violated if the transfer was affected by the individual's behaviour or environment.[8]

Finally, consider the intuition underlying this identification scheme. Given the nature of the endogeneity of education, we need to account for the relationship between u_i^* and the $v_i^{*j\prime}s$. Thus, consider two individuals with 'identical' parents (i.e. identical $v_i^{*j\prime}s$), but different socioeconomic backgrounds (i.e. $X_i's$). As these individuals are exposed to the same $v_i^{*j\prime}s$ they each have the same u_i^*. In the absence of heteroskedasticity the mapping of the $v_i^{*j\prime}s$ to the $u_i^{*\prime}s$ is the same as that of the $v_i^{j\prime}s$ to the $u_i's$ and the contribution of un observed ability to each individual educational level is the same. Thus there is no variation in the $X_i's$ which can be exploited to uncover the relationship between u_i^* and the $v_i^{*j\prime}s$. However, in the presence of heteroskedasticity the $v_i^{j\prime}s$, and thus the $u_i's$, will differ across the two individuals, and this will result in different education levels for both the parents and the children. These differences in education levels resulting from the heteroskedasticity provides the variation required to estimate the relationship between the u_i^* and the $v_i^{*j\prime}s$.

[6]Klein and Vella (2010) prove identification in the case where $Z = X$ noting that the choice of Z has no implications for what is a suitable instrument. In empirical applications it is likely that the conditional mean and variances may not be functions of the identical variables.

[7]The Monte Carlo evidence in Klein and Vella (2010) indicates the estimator works well even in the absence of parametric assumptions.

[8]Klein and Vella (2009, 2010) show that identification is also possible in the presence of more general error structures than that in equations (4) and (5). In Klein and Vella (2009) the error terms include an additional common unobserved component. That is, $u_i = H_u(X_i)\varpi_i u_i^*$ and $v_i^j = H_v^j(X_i)\varpi_i v_i^{*j}$. In our framework this additional component, ϖ_i, might be interpreted as unobserved factors that are common to parents and children and are not genetically transmitted (i.e. cultural traits or other unobservable characteristics of the economic and social context). Note that ϖ_i can also be interpreted as measurement error. Thus if the required identification conditions are satisfied our estimation strategy provides consistent estimates even in the presence of measurement error.

Before proceeding to the empirical section, it is useful to consider some possible concerns with our approach. First, our identifying restriction is that the transfer of unobserved traits is invariant to socioeconomic circumstances. Whereas this seems a reasonable assumption, we acknowledge that there are circumstances under which it may be possibly violated. However, our view is that the restriction is sensible and it is valuable to consider the estimates which are produced when it is imposed. This is also the case when alternative estimates of the intergenerational transmission process are obtained based on different identifying restrictions. Second, our approach imposes that the components of the control functions determining the mapping between the homoskedastic errors and the child's education level, denoted $\rho^j H_{ui}/H^j_{vi}$, depend only on the exogenous variables. Given the inherent circularity between the parent's education level and the $H^{j\prime}_{vi}s$ it is not possible to include this variable as an additional component in the control functions. To allow the control functions to vary by parental education, one could estimate the model separately for education groups. The relatively small sample sizes here do not allow us to estimate the model in such a manner. However, the relationship between the form of the control functions and the parental educational level is an issue worth investigating in future work. Finally, the form of the heteroskedastic functions requires that we impose a parametric relationship and a single index representation of how the $X's$ enter the model. The single index representation is required to make the procedure implementable and is a feature of the vast majority of econometric procedures. The sensitivity of the parametric form of the heteroskedasticity is an issue we address in the empirical section.

III. Results

We estimate the intergenerational transfer of education for a sample of individuals drawn from the National Longitudinal Survey of Youth 1979 (NLSY79). This survey comprises a representative sample of individuals living in the US aged 14–22 years in 1979. The survey was conducted annually until 1994 and every 2 years subsequently. While there are no variables which could be employed as plausible exclusions to estimate the model by instrumental variables, the parental information collected in 1979 allows estimation via the procedure discussed above.

The outcome on which we focus is years of schooling based on questions related to the highest grade of education completed. To reduce censoring of ongoing education activities we employ the information from the 1994 wave when the respondents are aged between 29 and 37 years and we assume they have completed their education. The highest grade of education completed by their parents is reported in the 1979 wave. The explanatory variables, aside from parent's education, are those typically employed in studies of education transmission and are listed with their summary statistics in Table 1. We restrict our analysis to the core sample of the NLSY79.[9] Following previous studies of intergenerational transmission we focus only on children raised in complete families based on whether the individual lived with both parents at the age of 14 years. We also exclude 23 individuals who report less than 8 years of completed education. The sample comprises 2,072 males and 2,282 females.

[9]The NLSY79 core subsample is constructed to be representative of the US population.

TABLE 1

Summary statistics

	All children	*Sons*	*Daughters*
Children's variables			
Years of education completed in 1994	13.38 (2.48)	13.30 (2.57)	13.45 (2.40)
Attended public school	0.93	0.93	0.93
Born in the US	0.95	0.95	0.95
Living in a Southern state at the age of 14	0.35	0.34	0.35
Living in a city at the age of 14	0.77	0.77	0.78
Gender (male = 1)	0.48		
Black	0.21	0.21	0.21
Hispanic	0.17	0.17	0.17
Non-black; non-Hispanic	0.62	0.62	0.62
Age in 2006	43.32 (2.19)	43.21 (2.21)	43.42 (2.17)
Parents' variables			
Mother's years of education	11.29 (3.09)	11.37 (3.06)	11.21 (3.11)
Father's years of education	11.34 (4.97)	11.43 (3.98)	11.26 (3.96)
Foreign born (mother)	0.10	0.09	0.10
Foreign born (father)	0.08	0.08	0.08
Mother's age in 1979	44.85 (6.66)	44.79 (6.74)	44.91 (6.58)
Father's age in 1979	48.05 (7.46)	47.91 (7.46)	48.18 (7.45)
Number of observations	4,354	2,072	2,282

Notes: Standard Deviations in parentheses.

Table 2 reports the estimates of the intergenerational transmission of education model.[10] The first column contains the OLS estimates and the coefficients on the education of each parent are statistically significant and indicate that for each year of father's education the individual acquires an additional 0.17 years while the corresponding effect for mother's education is 0.21 years. These estimates are consistent with the existing OLS results.

To employ the estimation strategy discussed above we require the residuals from the parents' education equations and estimates of the functions generating the conditional heteroskedasticity. Table 3 reports the estimates for the parent's equations. The effects are similar for both equations so we discuss them together. The negative age coefficients probably capture cohort effects and reflect the increasing level of education acquired by more recent birth cohorts. Being born overseas has a large negative and statistically significant effect on the educational attainment of both parents. To capture some regional and additional background characteristics we include the race of the child and indicators that the child was raised in a city and in the South. Whereas it is preferable to use the background variables of the parents this reduced the sample size and as there are no statistical difficulties introduced by employing these proxies this is the strategy we prefer. Note that the coefficients reflecting race effects show that parents of blacks and Hispanic children

[10]In choosing the explanatory variables that enter the empirical model we follow previous studies and include the age and gender of the child, as well as race and ethnic indicators (Plug, 2004; Black *et al.*, 2005). To better capture the presence of heteroskedasticity in the data we include a set of additional regressors: parental age, two geographical variables (i.e. living in a Southern state and in a city at the age 14) and a public school attendance indicator. Since some of these variables could be determined by parental education, we re-estimate the model excluding those additional regressors. Our main conclusions are unaffected and the results are available upon requests from the authors.

TABLE 2

Relationships between parents' and children's education

	OLS	*CF*
Mother's years of education	0.210 (0.015)	0.098 (0.037)
Father's years of education	0.167 (0.012)	0.021 (0.040)
Attended public school	−0.631 (0.133)	−0.602 (0.144)
Born in the US	0.283 (0.186)	0.270 (0.194)
Living in a Southern state at the age of 14	−0.001 (0.074)	−0.108 (0.080)
Living in a city at the age of 14	0.010 (0.081)	0.309 (0.101)
Mother's age	0.019 (0.009)	0.010 (0.009)
Father's age	0.025 (0.008)	0.015 (0.009)
Gender (male = 1)	−0.198 (0.066)	−0.202 (0.065)
Black	−0.041 (0.091)	−0.576 (0.122)
Hispanic	0.039 (0.113)	−0.917 (0.151)
Foreign born (mother)	0.696 (0.165)	0.464 (0.170)
Foreign born (father)	0.759 (0.175)	0.547 (0.200)
Child's age	−0.013 (0.016)	−0.010 (0.016)
Constant	7.943 (0.724)	11.72 (0.894)
ρ^M		0.100 (0.029)
ρ^F		0.177 (0.047)
Number of observations	4,354	4,354

Notes: Standard errors calculated from 1,000 bootstrap replications with random replacement.

TABLE 3

Parental education

	Mothers	*Fathers*
Years of education (mean) OLS		
Age in 1979	−0.018 (0.006)	−0.065 (0.007)
Foreign born	−1.303 (0.154)	−1.344 (0.214)
Black	−1.117 (0.108)	−2.618 (0.141)
Hispanic	−3.636 (0.124)	−3.827 (0.161)
Living in a Southern state at the age of 14	−0.417 (0.090)	−0.482 (0.118)
Living in a city at the age of 14	0.724 (0.098)	1.311 (0.129)
Constant	12.65 (0.290)	14.93 (0.365)
Breusch-Pagan test	1004.53	395
White test	638	421
Years of education (variance) NLLS		
Age in 1979	0.033 (0.009)	0.026 (0.005)
Foreign Born	0.859 (0.164)	0.540 (0.148)
Black	0.551 (0.191)	0.387 (0.110)
Hispanic	1.845 (0.171)	0.818 (0.107)
Living in a Southern state at the age of 14	−0.096 (0.176)	0.182 (0.091)
Living in a city at the age of 14	0.616 (0.314)	0.767 (0.129)
Constant	−2.361 (0.601)	−1.064 (0.331)
Number of observations	4,354	4,354

Notes: Standard errors calculated from 1,000 bootstrap replications with random replacement.

obtain significantly less education than those of whites. There are also differences by region and for those living in a city.

The test statistics for heteroskedasticity are also reported in Table 3 along with, in the lower panel, the estimates of the heteroskedastic functions for the parents' education equations.[11] Given the form we have assumed for H_i^{j2}, we can directly interpret the sign of these coefficients. Those for age and the immigrant indicator are both positive and statistically significant and reflect a higher variance in the schooling residuals for older and foreign born individuals. The residual variance is also bigger for minority groups and those living in cities. The coefficient for the age variable may reflect greater variance in the opportunities to acquire education for the older cohorts. The result capturing regional variation is consistent with previous studies (see, for example, Klein and Vella, 2009; Rummery, Vella and Verbeek, 1999). That is, the distribution of educational facilities may vary across urban and rural areas and this affects not only the mean but also the variance of schooling levels. The result associated with the Hispanic indicator captures the heterogeneous nature of the group which identifies itself as Hispanic.

We now return to the estimation of the child's education level while accounting for the endogeneity of the parent's education. As we estimate both the determinants of the conditional mean and conditional variance simultaneously it is necessary to specify the variables generating the heteroskedasticity. While we experimented with different choices, including one which contained all the variables in the conditional mean, we focus our discussion on our preferred specification with fewer variables.[12] Under this specification the index generating the heteroskedasticity includes dummies to capture regional differences as well as the child's race or ethnic origin to account for the heterogeneous nature of this group. We also include a gender dummy, an indicator for public school attendance and indicators for whether the parents were born in the US. The estimates of this form of heteroskedasticity are displayed in the first column of Table 4. The variance of the education residuals is higher for individuals living in cities and for those with a foreign born father. In contrast to the results for the parents, Table 4 indicates a lower residual variance for individuals in the minority groups.

The estimates of the conditional mean of education are in the second column of Table 2 under the heading CF. Before we focus on the effect of primary interest we highlight some other results. First, the estimates for the exogenous variables for the OLS and the CF procedures are generally similar. Both reveal a negative effect from public school on years of education. Also, after controlling for other influences, females obtain more years of schooling. There is also evidence that schooling levels among individuals with foreign born parents are higher than for those with native born parents.

Now focus on the estimates of primary interest. The CF estimates reveal a substantial reduction in the coefficients for the parents' education variables. For example, the father's education coefficient is reduced to 0.02 and is no longer statistically significant while the

[11] Note that the presence of heteroskedasticity in the parental education equations is enough to identify the coefficients in the main equation (6). Both the Breusch-Pagan and the White test indicate a substantial amount of heteroskedasticity in these equations.

[12] We do not report the results from these alternative specifications using different conditioning variables for the heteroskedasticity. However, they are qualitatively similar as those reported here. The primary differences were in the significance levels of the coefficients in the index generating the heteroskedasticity.

TABLE 4

Heteroskedastic index for the education of the child

	All children	*Sons*	*Daughters*
Attended public school	−0.234 (0.122)*	−0.349 (0.192)	−0.134 (0.173)
Living in a Southern state at the age of 14	−0.020 (0.087)	−0.143 (0.133)	0.029 (0.114)
Living in a city at the age of 14	0.228 (0.112)	0.476 (0.160)	0.089 (0.117)
Hispanic	−1.086 (0.163)	−1.238 (0.261)	−1.002 (0.247)
Black	−0.668 (0.112)	−0.726 (0.203)	−0.584 (0.144)
Foreign born (mother)	0.280 (0.188)	0.453 (0.295)	0.057 (0.229)
Foreign born (father)	0.411 (0.208)	0.545 (0.318)	0.483 (0.278)
Gender (male = 1)	0.036 (0.078)		
Constant	1.014 (0.206)	0.869 (0.249)	1.131 (0.282)
Number of observations	4,354	2,072	2,282

Notes: Standard errors calculated from 1,000 bootstrap replications with random replacement.

mother's education coefficient decreases to 0.10 while retaining statistical significance. This reflects that the OLS estimates are confounded by the endogeneity of the education variables. Equally interesting are the coefficients capturing the transfer of unobserved ability. The coefficients for the mother's and father's control functions, denoted ρ^j, are 0.10 and 0.18, respectively and each is highly statistically significant. This indicates that parental education is not exogenous to that of the child and that unobservables affecting education are positively correlated across generations. This is consistent with the existing evidence that the correlation between parents' and children's education partially reflects the transfer of unobserved ability. That is, the OLS estimate is substantially larger than those that control for ability transmission. Our results are also consistent with the recent IV studies which suggest the mother's educational level has the strongest impact.

Before examining how the transfer of education may vary by the gender of the child it is also interesting to consider the impact of the control functions on the variables capturing that the individual is black or Hispanic. While the OLS estimates surprisingly indicated that neither have a role in educational attainment, the CF estimates indicate that once the parents' ability is controlled each has a large negative impact. The ability bias confounding the OLS estimates is clearly masking the extent to which minority groups are being disadvantaged in the education process.[13]

Table 5 addresses gender differences in the intergenerational transmission of education mechanism. Column 1 reports the estimates for sons and reveals no statistically significant direct effects from the educational attainment of either the mother or the father. However, the coefficient on the control function for each of the parents is significant. In contrast, the results for daughters shown in column 2, are similar to those for the whole sample. Table 5 also reveals gender differences in other variables such as being born in the US, in the South or in a city.

Now assess the economic significance of our findings noting that our evidence is important for the ongoing debate on educational transmission as it is directly based on the feature of the data which is understood to be responsible for the endogeneity of parental education. While a strict interpretation of the individual coefficient estimates for the parental

[13]This is consistent with findings of Kane (1994) and Neal (2005).

TABLE 5

Relationships between parents' and children's education by gender of the child

	CF (sons)	*CF (daughters)*
Mother's years of education	0.070 (0.057)	0.130 (0.050)
Father's years of education	0.072 (0.058)	−0.006 (0.062)
Attended public school	−0.692 (0.230)	−0.561 (0.175)
Born in the US	−0.008 (0.329)	0.528 (0.252)
Living in a Southern state at the age of 14	−0.221 (0.117)	−0.006 (0.112)
Living in a city at the age of 14	0.420 (0.146)	0.182 (0.145)
Mother's age in 1979	0.006 (0.013)	0.014 (0.012)
Father's age in 1979	0.015 (0.013)	0.016 (0.011)
Black	−0.682 (0.192)	−0.397 (0.173)
Hispanic	−1.010 (0.230)	−0.752 (0.239)
Foreign born (mother)	0.237 (0.267)	0.677 (0.229)
Foreign born (father)	0.834 (0.330)	0.340 (0.262)
Child's age	−0.010 (0.024)	−0.011 (0.022)
Constant	11.73 (1.25)	11.18 (1.27)
ρ^M	0.128 (0.047)	0.080 (0.033)
ρ^F	0.138 (0.075)	0.178 (0.068)
Number of observations	2,072	2,282

Notes: Standard errors calculated from 1,000 bootstrap replications with random replacement.

education variables is that there is no effect from parents for sons and that there is only a mother's effect for daughters, an alternative interpretation is that the sum of the two parental education effects is equal for both genders. Such an interpretation would be consistent with the presence of positive sorting in the marital market where parental education levels are highly correlated. Accordingly direct education effects might exist for sons but the high correlation between the parents' education makes it difficult to disentangle the individual contribution from each parent. This, in fact, is supported by the data and our results. The correlation between father's and mother's education is 0.78. Moreover, while for sons both parents education levels are individually statistically insignificant the null hypothesis that they are jointly zero is rejected with a t-statistic of 2.91. The evidence regarding the role of unobserved ability is far clearer. The transfer of unobserved ability from both parents has a statistically significant and large positive effect on the education level of the sons. Moreover, the coefficients are approximately equal.

The evidence for daughters portrays a somewhat different story. First, the education coefficients strongly suggest a direct effect from the mother's education while there is no effect from that of the father. Note, however, that the sum of the coefficients for the mother and the father is approximately equal for sons and daughters. This indicates that in the case where the parents have the same educational levels the contribution of parental education is the same for daughters and sons. Second, for daughters we are able to disentangle the direct effect of education from that of unobserved ability. That is, we find a statistically significant role for both the mother's education and her unobserved ability transfer. Finally, the transfer from fathers to daughters is only through the unobserved ability component.

In addition to supporting the earlier evidence that the transfer of unobserved ability is confounding the OLS estimates the most interesting finding of this article is the difference in the results for sons and daughters. While there is a remarkable symmetry in the role of parents for sons this symmetry is absent for daughters. That is, mothers and fathers play quite different roles for their daughters. While daughters benefit a great deal from the transfer of unobserved ability from their father, the mother's educational behaviour, in addition to her ability transfer, is of consequence to the daughter's education.

One possible explanation of the result for mothers and daughters may be related to the positive correlation in the economic behaviour of women across successive generations, which has been partially attributed to the intergenerational transfer of cultural traits (see, for example, Farré and Vella, 2007; Fernández, 2007). While it is difficult to identify the determinants of this intergenerational correlation, it does seem to manifest itself in our data through a strong correlation in educational attainment. Our results, in conjunction with this related literature, suggest that this correlation not only reflects 'role model' effects but also some effect due to the transfer of unobserved ability from mothers to daughters.

IV. Conclusion

Our evidence strongly suggests that the OLS estimates of the intergenerational transmission of education are biased upwards because of the transfer of unobserved ability and that the bias is large. For both sons and daughters we find that the inherited endowment of unobserved ability, from both parents, is an important determinant of their education. The coefficients capturing this transmission mechanism are large although for daughters the impact of the father's ability is larger than that of the mother. This might reflect that a mother's influence on her daughter's behaviour is shared over both her education and her transfer of ability. For the effect of parental education levels we conclude that both for daughters and sons there are intergenerational effects and they appear to be of the same magnitude. However, for daughters the effects are attributed to the mother while for sons we are unable to distinguish whether they are due to the father or mother. We conclude that the high correlation between parents' education and the important role model mothers play for their daughters are responsible for this result.

Final Manuscript Received: June 2011

References

Antonovics, K. L. and Goldberger, A. S. (2005). 'Does increasing women's schooling raise the schooling of the next generation? Comment', *American Economic Review*, Vol. 95, pp. 1738–1744.

Behrman, J. R. (1997). *Mother's Schooling and Child Education: A Survey*, Penn Institute For Economic Research Working Papers 97(025), University of Pennsylvania.

Behrman, J. R. and Rosenzweig, M. R. (2002). 'Does increasing women's schooling raise the schooling of the next generation', *American Economic Review*, Vol. 95, pp. 323–334.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005). 'Why the apple doesn't fall far: understanding intergenerational transmission of human capital', *American Economic Review*, Vol. 95, pp. 437–449.

Björklund, A., Lindahl, M. and Plug, E. (2006). 'The origins of intergenerational associations: lessons from Swedish adoption data', *The Quarterly Journal of Economics*, Vol. 121, pp. 999–1028.

Chevalier, A. (2004). *Parental Education and Child's Education: A Natural Experiment*, IZA Discussion Papers No. 1153.

Farré, L. and Vella, F. (2007). *The Intergenerational Transmission of Gender Role Attitudes and Its Implications for Female Labor Force Participation*, IZA Discussion Papers No. 2802.
Fernandez, R. (2007). 'Women, work, and culture', *Journal of the European Economic Association*, Vol. 5, pp. 305–332.
Haveman, R. and Wolfe, B. (1995). 'The determinants of children attainments: a review of methods and findings', *Journal of Economic Literature*, Vol. 33, pp. 1829–1878.
Holmlund, H., Lindahl, M. and Plug, E. (2011). 'The causal effect of parents' schooling on children's schooling: a comparison of estimation methods', *Journal of Economic Literature*, forthcoming.
Kane, T. J. (1994). 'College entry by blacks since 1970: the role of college costs, family background, and the returns to education', *The Journal of Political Economy*, Vol. 103, pp. 878–911.
Klein, R. and Vella, F. (2009). 'Estimating the return to endogenous schooling decisions for Australian workers via conditional second moment', *Journal of Human Resources*, Vol. 44, pp. 1047–1065.
Klein, R. and Vella, F. (2010). 'Estimating a class of triangular simultaneous equations models without exclusion restrictions', *Journal of Econometrics*, Vol. 154, pp. 154–164.
Neal, D. (2005). 'Why has black-white skill convergence stopped?', NBER Working Papers 11090.
Plug, E. (2004). 'Estimating the effect of mother's schooling on children's schooling using a sample of adoptees', *American Economic Review*, Vol. 92, pp. 358–368.
Rummery, S., Vella, F. and Verbeek, M. (1999). 'Estimating the returns to education for Australian youth via rank-order instrumental variables', *Labour Economics*, Vol. 6, pp. 777–792.
Sacerdote, B. (2004). *What Happens When We Randomly Assign Children to Families?* NBER Working Paper 10894.

Appendix

Here we outline the logic underlying the Klein and Vella (2010) procedure (hereafter KV) but refer the reader to the KV paper for formal proofs. The model is:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + u_i, \quad i = 1, \ldots, N \tag{1A}$$

$$S_i^j = X_i\delta_j + v_i^j, \quad j = M, F \tag{2A}$$

where the correlation of the error terms across equations renders the OLS estimates of equation (1A) inconsistent. Consider the control function version of instrumental variables estimation for this model. This requires purging equation (1A) of the component of the error term which is correlated with the reduced form errors. That is, recall that the main equation error can be written:

$$u_i = \lambda^M v_i^M + \lambda^F v_i^F + e_i \tag{3A}$$

where $\lambda^j = \text{cov}(v^j u)/\text{var}(v^j)$ when there is no dependence between the error distributions and the $X's$. This procedure requires estimates of the two reduced forms errors which can then be used to estimate:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + \lambda^M \hat{v}_i^M + \lambda^F \hat{v}_i^F + e_{1i}, \tag{4A}$$

where the e_{1i} represents a zero mean error term. Estimation of equation (4A) is not possible however as the absence of exclusion restrictions in the reduced form equations ensures the matrix $M = [X, S, S^M, S^F, \hat{v}^M, \hat{v}^F]$ is not of full rank.

Now assume the errors distributions depend on the $X's$ (e.g. heteroskedasticity). The coefficients in equation (3A) become:

$$\lambda^j = \frac{\mathrm{cov}(v^j u \mid X)}{\mathrm{var}(v^j \mid X)} = A^j(X) \tag{5A}$$

which implies that the impact of v_i^j on u_i depends on the value of X_i. Under the conditional correlation assumption KV show:

$$\lambda^j = \frac{\mathrm{cov}(v^j u \mid X)}{\mathrm{var}(v^j \mid X)} = \rho^j \frac{\sqrt{\mathrm{Var}(u_i \mid X_i)}}{\sqrt{\mathrm{Var}(v_i^j \mid X_i)}} \tag{6A}$$

which given our assumptions in the text gives:

$$u_i = \rho^M \frac{H_{ui}}{H_{vi}^M} v_i^M + \rho^F \frac{H_{ui}}{H_{vi}^F} v_i^F + e_{1i} \tag{7A}$$

Estimation is now feasible as the matrix $M^1 = [X, S, S^M, S^F, \frac{H_u}{H_v^M}\hat{v}^M, \frac{H_u}{H_v^F}\hat{v}^F]$ is of full rank due to the nonlinearity induced by the multiplicative role of the $X's$. KV show that the parameters of the model are identified even without parametric assumptions regarding the form of heteroskedasticity.

While KV (2010) provide a semiparametric estimation procedure for equation (4A), retaining the semiparametric aspect in practice is associated with demanding computational requirements. Thus we employ the following parametric version:

(i) Regress S^M and S^F on X to get $\hat{v}^M$ and $\hat{v}^F$.

(ii) Use assumption (8) and estimate θ_{2j} through nonlinear least squares using $\ln(\hat{v}^{j2})$ as the dependent variable. With these estimates compute $\hat{H}_{vi}^j = \sqrt{\exp(Z_i^j \hat{\theta}_2)}$.

(iii) To estimate the primary equation parameters we can proceed in two ways. First, given our assumptions regarding the form for H_u we can estimate the parameters via the following nonlinear least squares problem:

$$\min_{\beta,\rho^j,\alpha_1,\theta_1} \sum_{i=1}^{N} \begin{pmatrix} S_i^C - X_i\beta_0 - \beta_M S_i^M - \beta_F S_i^F - \rho^M \left(\sqrt{\exp(Z_{ui}\theta_1)}\right) * \frac{\hat{v}_i^M}{\hat{H}_{vi}^M} \\ -\rho^F \left(\sqrt{\exp(Z_{ui}\theta_1)}\right) \frac{\hat{v}_i^F}{\hat{H}_{vi}^F} \end{pmatrix}^2 . \tag{8A}$$

While this produces consistent estimates of the unknown parameters in equation (8A) it requires the estimation of H_u through the minimization of a least squares problem related to S^C. An alternative approach is to estimate θ_1 in H_u in a similar manner as for the parental education equations. Accordingly for a given value of β, say β_c, we define the residual $u(\beta_c)$. Using this value of $u(\beta_c)$, we regress $u(\beta_c)^2$ on $Z_{ui}\theta_{cu}$ where we also use a candidate value for θ_{cu}. From this regression, we compute $\hat{H}_u(\beta_c)$ as $\sqrt{(Z_{ui}\theta_{cu})}$ and estimate the $\rho's$ as:

$$\min_{\rho_c^j} \sum_{i=1}^{N} \left(u_i(\beta_c) - \rho_c^M \frac{\hat{H}_{ui}(\beta_c)}{\hat{H}_{vi}^M} \hat{v}_i^M - \rho_c^F \frac{\hat{H}_{ui}(\beta_c)}{\hat{H}_{vi}^F} \hat{v}_i^F \right)^2 . \tag{9A}$$

Consistent estimates of the unknown parameters in equation (9A) are obtained by searching over β_c, θ_{cu} and ρ_c^j. With these estimates of β, which we denote β_f, we define the residual $u_{if} = S_i^C - X_i\beta_{f0} - \beta_{fM} S_i^M - \beta_{fF} S_i^F$. We then use u_{if}^2 to get $\hat{H}_u(\beta_f)$ in precisely the same way as in step (ii). With $\hat{H}_u(\beta_f)$ we then regress S_i^C on X_i, S_i^j and $(\hat{H}_u(\beta_f)/\hat{H}_{vi}^j\hat{v}_i^j)$ to get the final estimates. This final step separates the estimation of the $\beta' s$ from the estimation of H_u. Note, however, that in this particular example it gave almost identical estimates.