

IZA DP No. 3967

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

Lidia Farré  
Roger Klein  
Francis Vella

January 2009

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

**Lidia Farré**

*University of Alicante*

**Roger Klein**

*Rutgers University*

**Francis Vella**

*Georgetown University  
and IZA*

Discussion Paper No. 3967  
January 2009

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Does Increasing Parents' Schooling Raise the Schooling of the Next Generation? Evidence Based on Conditional Second Moments<sup>\*</sup>**

This paper 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 this causal effect. We conclude the observed intergenerational educational correlation reflects both a causal parental educational effect and a transfer of unobserved ability.

JEL Classification: C31, J62

Keywords: intergenerational mobility, endogeneity, conditional correlation

Corresponding author:

Francis Vella  
Department of Economics  
Georgetown University  
580 ICC, 37th & O Streets, NW  
Washington DC 20057  
USA  
E-mail: [fgv@georgetown.edu](mailto:fgv@georgetown.edu)

---

<sup>\*</sup> Farré thanks the Spanish Ministry of Education (Grant SEJ 2005-02829/ECON) for financial support.

# 1 Introduction

Although it is well established that a positive correlation exists between an individual's educational attainment and that of his/her parents it is unclear what it precisely captures.<sup>1</sup> While 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.

To identify this causal component some studies have focussed 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

---

<sup>1</sup>Haveman and Wolfe (1995) and Behrman (1997) provide extensive surveys of the earlier literature on the intergenerational transmission of education.

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 mother's education but no significant effect from the father's. This range of conclusions reflects the use of different data sets but highlights that alternative approaches may not identify the causal effect from the same part of the educational distribution. Holmlund, Lindahl and Plug (2008) 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.

While 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 behavior 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' socioeconomic 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 3 presents the data and our empirical results and also provides some concluding comments.

## 2 Empirical Model

Consider the following model of educational transfer:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + u_i, \quad i = 1, \dots, N \quad (1)$$

$$S_i^j = X_i\delta_j + v_i^j, \quad j = M, F. \quad (2)$$

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  $\beta$  inconsistent. As we allow the same  $X$  to enter (1) and (2) there is no exogenous source of variation in parents' education which identifies  $\beta$ . That is, there are no available instruments.

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

$$E[u_i|X_i] = E[v_i^j|X_i] = 0. \quad (3)$$

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:

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

where  $u_i^*$  and  $v_i^{*j}$  are correlated homoskedastic error terms which we interpret as measures of unobserved ability. According to (4) individuals receive values of  $u_i^*$  and  $v_i^{*j}$ , but the contribution of this unobserved ability to their educational achievement will depend on their respective socioeconomic environments or observed characteristics as determined by the relevant  $H$  function.

An implication of (4) is that the intergenerational transmission of unobserved ability operates through the relationship between  $u_i^*$  and the  $v_i^{*j'}$ s and not, necessarily, that between  $u_i$  and the  $v_i^{j'}$ 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.

OLS estimation of (1) produces inconsistent estimates due to the lack of orthogonality between the  $S_i^{j'}$ s and  $u_i$  and the moments corresponding to (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 correlations between the homoskedastic error terms are constant.<sup>2</sup> That is:

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

Following Klein and Vella (2006) these "constant conditional correlation coefficient"

---

<sup>2</sup>Klein and Vella (2006) 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.

moments can in conjunction with (3), and in the presence of (4), identify the model.<sup>3</sup>

Using these moments one can estimate the model by GMM. However, the estimation of these conditional moments is complicated due to their dependence on the unknown conditional variances and covariances. Klein and Vella (2006) show that the same moments can be imposed by estimating the following control function model:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + \rho^M \frac{H_{ui}}{\widehat{H}_{vi}^M} \widehat{v}_i^M + \rho^F \frac{H_{ui}}{\widehat{H}_{vi}^F} \widehat{v}_i^F + e_i \quad (6)$$

where  $\widehat{v}_i^M$  and  $\widehat{v}_i^F$  are the residuals from the parent's education equations;  $H_{ui}$  denotes the unknown  $H_u(X_i)$  while  $\widehat{H}_{vi}^j$  are the estimates of  $H_v^j(X_i)$ ; and  $e_i$  is a zero mean disturbance which is uncorrelated with the included regressors.<sup>4</sup>

Estimation of (6) is considerably simpler than the corresponding GMM procedure but is infeasible here due to the large dimension of  $X$  and the unknown nature of the  $H$  functions. Klein and Vella (2006) identify the parameters in (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 nonparametric and semiparametric representations of the heteroskedasticity. While this is theoretically attractive, as identification is not reliant on specific forms of heteroskedasticity, it is computationally demanding. However to reduce computation the  $H$  functions can be parameterized. Accordingly we specify the following form:

---

<sup>3</sup>Klein and Vella (2006) 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.

<sup>4</sup>Klein and Vella (2006) 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.



$$H_{ui}^2 = \exp(\alpha_1(Z_{ui}\theta_1)) \quad (7)$$

$$H_{vi}^{j2} = \exp(\alpha_2^j(Z_{vi}^j\theta_{2j})), \quad j = M, F \quad (8)$$

where the  $Z$ 's are the vector of variables considered to be responsible for the heteroskedasticity in the respective equations and  $\alpha$  and  $\theta$  are unknown parameters to be estimated.<sup>5</sup> One can also experiment with alternative functional forms for the heteroskedasticity. For example, below we employed a specification in which the  $H$ 's also included a quadratic term for the heteroskedastic index and found our main results were unaffected by this alternative specification.

Before proceeding consider whether the key 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.

---

<sup>5</sup>Klein and Vella (2006) 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.

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  $v_i^{*j'}$ 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 color, 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 behavior or environment.

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  $v_i^{*j'}$ s. Thus, consider two individuals with "identical" parents (i.e. identical  $v_i^{*j'}$ s), but different socioeconomic backgrounds (i.e.  $X_i'$ s). As these individuals are exposed to the same  $v_i^{*j'}$ s they each have the same  $u_i^*$ . In the absence of heteroskedasticity the mapping of the  $v_i^{*j'}$ s to the  $u_i^{*j'}$ s is the same as that of the  $v_i^{j'}$ s to the  $u_i^j$ s and the contribution of unobserved 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  $v_i^{*j'}$ s. However, in the presence of heteroskedasticity the  $v_i^{j'}$ s, and thus the  $u_i^j$ 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  $v_i^{*j'}$ s.

### 3 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 to 22 years in 1979. The survey was conducted annually until 1994 and every two 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 independent 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.<sup>6</sup> 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 2072 males and 2282 females.

Table 2 reports the estimates of the intergenerational transmission of education model. 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

---

<sup>6</sup>The NLSY79 core subsample is constructed to be representative of the US population.

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 parent's education equations and estimates of the functions generating the conditional heteroskedasticity.<sup>7</sup> 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. While 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 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 and the underlying index for the parents' education equations. Given the form we have assumed for  $H_i^{j2}$ , and the estimated positive coefficients on the index  $\alpha_2^j$ , 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

---

<sup>7</sup>The appendix provides a detailed discussion of how the estimator is implemented.

older and foreign born individuals. The residual variance is also bigger for minority groups and those living in cities. While we do not focus on the magnitude of these coefficients they appear reasonable.

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.<sup>8</sup> 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 heterogenous nature of this group. We also include a gender dummy 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 5. 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 5 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 completed years of education. Also, after controlling for

---

<sup>8</sup>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.

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 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  $\rho^j$ , 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.<sup>9</sup>

Table 4 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 4 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 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

---

<sup>9</sup>This is consistent with findings of Kane (1994) and Neal (2005).

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 paper 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 behavior, in addition to her ability transfer, is of consequence to the daughter's educational attainment.<sup>10</sup>

One possible explanation as to why we can identify a clear effect from both chan-

---

<sup>10</sup>Note, however, that the contribution of the two effects is approximately the same for both parents.



nels for mothers and daughters is that the mother's education might capture the existence of other factors which are also transferred across generations. For example, Farré and Vella (2007) provide evidence, using the same data examined here, that a daughter's attitude towards the role of women in the labor market is strongly correlated with that of her mother. This evidence suggests that the similarity in the economic behavior of females across generations may go beyond the impact of ability transfer and that mothers serve as important role models for their daughters. This is supported by Fernandez (2007) who finds that the work behavior of second-generation American women is similar to that of women in the country from which their parents migrate.

In conclusion our evidence strongly supports that the OLS estimates of the intergenerational transmission of education are biased upwards due to 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 educational attainment. The coefficients capturing this transmission mechanism are large although for daughters the impact of the father's unobserved ability is larger than that of the mother. This might reflect that a mother's influence on her daughter's behavior is shared over both her educational attainment and her transfer of unobserved ability. Regarding the direct 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.

## References

- [1] Antonovics, Kate L. and Arthur S. Goldberger. 2005. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Comment." *American Economic Review*, 95(5): 1738-44.
- [2] Behrman, Jere R. 1997. "Mother's Schooling and Child Education: A Survey." Penn Institute For Economic Research Working Papers 97(025). University of Pennsylvania.
- [3] Behrman, Jere R. and Mark R. Rosenzweig. 2002. "Does increasing Women's Schooling Raise the Schooling of the Next Generation." *American Economic Review*, 95(1): 323-334.
- [4] Black, Sandra E.; Paul J. Devereux and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review*, 95(1): 437-49.
- [5] Björklund, Anders; Mikael Lindahl and Erik Plug. 2006. "The origins of intergenerational associations: Lessons from Swedish adoption data." *The Quarterly Journal of Economics*, 121(3): 999-1028.
- [6] Chevalier, Arnaud. 2004. "Parental Education and Child's Education: A Natural experiment." IZA Discussion Papers No. 1153.
- [7] Farré, Lúdia and Francis Vella. 2007. "The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labor Force Participation." IZA Discussion Papers No. 2802.
- [8] Fernandez, Raquel. 2007. "Women, work, and culture." *Journal of the European Economic Association*, 5(2-3): 305-332.
- [9] Haveman, Robert and Barbara Wolfe. 1995. "The Determinants of Children Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33(4): 1829-78.
- [10] Holmlund, Helena; Mikael Lindahl and Erik Plug. 2006. "Estimating Intergenerational Schooling Effects: A Comparison of Methods." Mimeo.
- [11] Kane, Thomas 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*, 103(5): 878-911.

- [12] Klein, Roger and Francis Vella. 2006. "Estimating a Class of Triangular Simultaneous Equations Models Without Exclusion Restrictions." IZA Working Paper 2378.
- [13] Neal, Derek. 2005. "Why has black-white skill convergence stopped?." NBER Working Papers 11090.
- [14] Plug, Erik. 2004. "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees." *American Economic Review*, 92(2): 358-368.
- [15] Sacerdote, Bruce. 2004. "What happens when we randomly assign children to families?" NBER Working Paper 10894.

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	4354	2072	2282

\*Standard Deviations in parentheses

Table 2: Relationships between Parents' and Children's Education

	OLS	CF*
Mother's years of education	0.210 (0.015)	0.098 (0.036)
Father's years of education	0.167 (0.012)	0.021 (0.040)
Attended public school	-0.631 (0.133)	-0.602 (0.146)
Born in the US	0.283 (0.186)	0.270 (0.198)
Living in a Southern state at the age of 14	-0.001 (0.074)	-0.108 (0.082)
Living in a city at the age of 14	0.010 (0.081)	0.309 (0.106)
Mother's age	0.019 (0.009)	0.010 (0.008)
Father's age	0.025 (0.008)	0.015 (0.009)
Gender (male=1)	-0.198 (0.066)	-0.202 (0.066)
Black	-0.041 (0.091)	-0.576 (0.125)
Hispanic	0.039 (0.113)	-0.917 (0.149)
Foreign born (mother)	0.696 (0.165)	0.464 (0.164)
Foreign born (father)	0.759 (0.175)	0.547 (0.200)
Child's age	-0.013 (0.016)	-0.010 (0.017)
Constant	7.943 (0.724)	11.72 (0.939)
$\rho^M$		0.100 (0.029)
$\rho^F$		0.177 (0.047)
Number of observations	4354	4354

\*Standard errors calculated from 500 bootstrap replications with random replacement

Table 3: Parental Education

	Mothers	Fathers
<b>Years of education (mean) OLS</b>		
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
<b>Years of education (variance) NLLS*</b>		
Age in 1979	0.034 (0.009)	0.029 (0.006)
Foreign Born	0.883 (0.169)	0.595 (0.170)
Black	0.566 (0.196)	0.426 (0.123)
Hispanic	1.897 (0.169)	0.900 (0.116)
Living in a Southern state at the age of 14	-0.098 (0.190)	0.200 (0.109)
Living in a city at the age of 14	0.634 (0.327)	0.844 (0.130)
Constant	-2.428 (0.586)	-1.171 (0.331)
Index coefficient ( $\alpha_2$ )	0.973 (0.012)	0.909 (0.016)
Number of observations	4354	4354

\*Standard errors calculated from 500 bootstrap replications with random replacement

Table 4: Relationships between Parents' and Children's Education by gender of the child\*

	CF (Sons)	CF (Daughters)
Mother's years of education	0.069 (0.054)	0.129 (0.051)
Father's years of education	0.069 (0.058)	-0.006 (0.063)
Attended public school	-0.694 (0.228)	-0.561 (0.175)
Born in the US	-0.009 (0.327)	0.528 (0.262)
Living in a Southern state at the age of 14	-0.222 (0.114)	-0.006 (0.109)
Living in a city at the age of 14	0.426 (0.150)	0.182 (0.150)
Mother's age in 1979	0.006 (0.013)	0.014 (0.019)
Father's age in 1979	0.015 (0.013)	0.016 (0.011)
Black	-0.693 (0.193)	-0.398 (0.176)
Hispanic	-1.028 (0.222)	-0.752 (0.234)
Foreign born (mother)	0.236 (0.266)	0.677 (0.239)
Foreign born (father)	0.831 (0.323)	0.339 (0.264)
Child's age	-0.010 (0.024)	-0.011 (0.022)
Constant	11.80 (1.25)	11.17 (1.25)
$\rho^M$	0.130 (0.044)	0.080 (0.036)
$\rho^F$	0.143 (0.073)	0.179 (0.074)
Number of observations	2072	2282

\*Standard errors calculated from 500 bootstrap replications with random replacement



Table 5: Heteroskedastic index for the Education of the Child\*

	All Children	Sons	Daughters
Attended public school	-0.256 (0.133)	-0.354 (0.197)	-0.061 (0.192)
Living in a Southern state at the age of 14	-0.022 (0.093)	-0.163 (0.144)	0.015 (0.123)
Living in a city at the age of 14	0.249 (0.126)	0.496 (0.176)	0.078 (0.137)
Hispanic	-1.186 (0.172)	-1.256 (0.273)	-1.001 (0.243)
Black	-0.729 (0.119)	-0.810 (0.209)	-0.626 (0.150)
Foreign born (mother)	0.306 (0.200)	0.482 (0.306)	0.189 (0.256)
Foreign born (father)	0.449 (0.223)	0.546 (0.321)	0.271 (0.281)
Gender (male=1)	0.039 (0.088)		
Constant	0.915 (0.219)	0.917 (0.251)	1.238 (0.282)
Index coefficient ( $\alpha_1$ )	1.107 (0.014)	0.932 (0.019)	0.907 (0.021)
Number of observations	4354	2072	2282

\*Standard errors calculated from 500 bootstrap replications with random replacement

## 4 Appendix

Here we outline the logic underlying the Klein and Vella (2006) procedure (hereafter KV) but refer the reader to the KV paper for formal proofs. The model is:

$$S_i^C = X_i\beta_0 + \beta_M S_i^M + \beta_F S_i^F + u_i, \quad i = 1, \dots, N \quad (1A)$$

$$S_i^j = X_i\delta_j + v_i^j, \quad j = M, F \quad (2A)$$

where the correlation of the error terms across equations renders the OLS estimates of (1A) inconsistent. Consider the control function version of instrumental variables estimation for this model. This requires purging (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:

$$u_i = \lambda^M v_i^M + \lambda^F v_i^F + e_i \quad (3A)$$

where  $\lambda^j = \frac{\text{cov}(v^j, u)}{\text{var}(v^j)}$  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:

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

where the  $e_{1i}$  represents a zero mean error term. Estimation of (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 (3A) become:

$$\lambda^j = \frac{\text{cov}(v^j u | X)}{\text{var}(v^j | X)} = A^j(X)$$

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:

$$\lambda^j = \frac{\text{cov}(v^j u | X)}{\text{var}(v^j | X)} = \rho^j \frac{\sqrt{\text{Var}(u_i | X_i)}}{\sqrt{\text{Var}(v_i^j | X_i)}}$$

which given our assumptions in the text gives:

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

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 non linearity 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 (2006) provide a semiparametric estimation procedure for (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  $\alpha_2^j$  and  $\theta_{2j}$  through non linear least squares using  $\ln(\hat{v}^{j2})$  as the dependent variable. With these estimates compute  $\hat{H}_{vi}^j = \sqrt{\exp(\hat{\alpha}_2^j(Z_i^j \hat{\theta}_2))}$ .
- 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 non linear least squares problem:

$$\min_{\beta, \rho^j, \alpha_1, \theta_1} \sum_{i=1}^N \left( \begin{array}{c} S_i^C - X_i \beta_0 - \beta_M S_i^M - \beta_F S_i^F - \rho^M \left( \sqrt{\exp \alpha_1(Z_{ui} \theta_1)} \right) * \frac{\hat{v}_i^M}{\hat{H}_{vi}^M} \\ - \rho^F \left( \sqrt{\exp \alpha_1(Z_{ui} \theta_1)} \right) \frac{\hat{v}_i^F}{\hat{H}_{vi}^F} \end{array} \right)^2. \quad (5A)$$

While this produces consistent estimates of the unknown parameters in (5A) 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  $\alpha_1$  and  $\theta_1$  in  $H_u$  in a similar manner as for the parental education equations. Accordingly for a given value of  $\beta$ , say  $\beta_c$ , 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  $\theta_{cu}$ . From this regression we compute  $\hat{H}_u(\beta_c)$  as  $\sqrt{\hat{\alpha}_1(Z_{ui} \theta_{cu})}$  and estimate the  $\rho$ 's as:

$$\min_{\rho_c^j} \sum_{i=1}^N u_i(\beta_c) - \rho_c^M \frac{\hat{H}_{ui}(\beta_c)}{\hat{H}_{vi}^M} \hat{v}_i^M - \rho_c^F \frac{\hat{H}_{ui}(\beta_c)}{\hat{H}_{vi}^F} \hat{v}_i^F. \quad (6A)$$

Consistent estimates of the unknown parameters in (6A) are obtained by searching over  $\beta_c$ ,  $\theta_{cu}$  and  $\rho_c^j$ . With these estimates of  $\beta$ , which we denote  $\beta_f$ , 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  $\frac{\hat{H}_u(\beta_f)}{\hat{H}_{vi}^j} \hat{v}_i^j$  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.