

An examination of paternal and maternal intergenerational transmission of schooling

Chiara Pronzato

Received: 4 November 2008 / Accepted: 18 January 2010 /
Published online: 5 March 2010
© Springer-Verlag 2010

Abstract More educated parents are observed to have better educated children. However, previous research has found conflicting results regarding the role of fathers and mothers: in most cases, a strong positive paternal effect was found with a negligible maternal effect; in fewer cases, opposite results were found. In this paper, I use a sample of Norwegian twins to evaluate the impact of sample size and sample selection on the estimates' robustness: results concerning the effect of mother's education are very sensitive to the sample size, while the selection of the sample seems to be a key to reconciling previous results.

Keywords Intergenerational transmission · Education · Twin-estimator

JEL Classification C23 · I2

Responsible editor: Deborah Cobb-Clark

C. Pronzato (✉)
DONDENA Centre for Research on Social Dynamics, Bocconi University, Via Röntgen n. 1,
20136, Milan, Italy
e-mail: chiara.pronzato@unibocconi.it

C. Pronzato
Statistics Norway, Oslo, Norway

C. Pronzato
CHILD Collegio Carlo Alberto, Turin, Italy

C. Pronzato
ISER University of Essex, Colchester, UK

1 Introduction

Much research, in recent years, has focused on the link between parental education and children's education. It has been empirically shown that more educated parents have better educated children. Using simple regression analysis, the correlation between parent's and child's education is strong and robust to a number of controls, sample selections, and countries (Haveman and Wolfe 1995; Hertz et al. 2007).

The policy implications of a causal link between parental education and children's education are huge. Increasing education today would lead to an increase in the schooling of the next generation and, in this way, to an improvement of later life outcomes such as health, productivity, and wealth. From a policy point of view, however, it is important to distinguish between causation and selection. Better educated parents have, for example, higher ability, which partially transmits to their children. Better schooling performances of their children could be simply the outcome of this genetic transmission.

When researchers have tried to control for ability and other unobserved characteristics of the parental environment, they have found conflicting results. In most cases, they have found a strong positive father's effect with a negligible mother's effect. In a few cases, a positive effect has been found for the mother and not for the father. The results depend on different identification strategies and on different sources of information. In Section 2, techniques and results are summarized from previous studies, with particular attention paid to the work by Holmlund et al. (2008). Their study uses different identification strategies with the same data source, concluding that the identification strategy matters. My paper is in the same spirit as Holmlund et al. (2008), trying to shed some light on the intergenerational transmission of schooling by repeating some previous analyses with new and rich data and by testing what may drive the results. While Holmlund et al. (2008) employ different identification strategies using the same data, I only make use of one strategy, the twin-estimator, and focus on how sample size and heterogeneous effects in the population may help explain different results found in the literature.

To this aim, I use Norwegian register data, a very rich source of information, which provides demographic, educational, and income characteristics for the whole population for the years 1993–2001 (Section 3). In Section 4, the twin-estimator is employed to identify the intergenerational transmission of schooling, and further analyses are carried out to check how robust estimates are when varying the sample size and when selecting different parts of the population. In Section 5, the analysis is repeated by using siblings instead of twins. The size of the bias in the estimates, which we may get from the study of siblings, is informative on the nature of what we call “family unobservables.” This section also provides a benchmark for analyses carried out with siblings, which are more easily available in survey data. Conclusions follow in Section 6.

The paper finds a positive and strong effect of father's education on children's education, as in previous papers, but also a positive and significant effect of mother's education, even if smaller in absolute value. Both sample size and

heterogeneous effects along the educational distribution are shown to be important for explaining different results by gender. The paper shows that, with relatively small samples (typical of twins studies), we are always more likely to find a significant effect of father's education than of mother's. By focusing on different parts of the distribution of parents' education, the paper finds that the effect of father's education is very strong but only at the top part of the distribution, while 1 year more of mother's education seems to matter more for low educated mothers. Results obtained with siblings go in the same direction, but tend to underestimate the effect of one additional year of father's schooling.

2 The intergenerational transmission of education in the economic literature

There are typically three strategies used in the literature to identify the effect of parental education on children's education: identical twins, adopted children, and an instrumental variable approach where the instrument is a reform of the educational system. All strategies aim to separate the effect of parental education from the effect of other unobservable characteristics (e.g., cognitive and noncognitive skills) that are correlated with parental education, are transmitted from parents to children, and influence children's schooling attainment. They will be generally called family unobservables throughout the paper.

This paper uses monozygotic and dizygotic twins to determine the effect of parental education on children's education (see Sections 3 and 4). In fact, twins are the most similar individuals we can observe: they share the same family background, they experience lifetime events at the same time, and they share the same genes (only if monozygotic twins). When studying the intergenerational transmission of education, we compare the schooling of twins' children (i.e., cousins). The cousins share to the same degree the ability and other family features transmitted by the twins. On the other hand, they are exposed to different treatments: they can be the children of the more educated twin, the other parent (twin's spouse) has different characteristics, and some of them can grow up in a single-parent household. In such studies, we can identify the effect of parental education on children's education, looking whether the child of the more educated twin has higher educational attainment than the child of the less educated twin. The shortcoming of this strategy is the assumption of random education between twins: why do children with identical abilities and experiences end up with different levels of education? If there are some characteristics which make one twin gain more education than the other, and if these characteristics can be transmitted to their children, then the resulting estimates are still biased. Despite this strong assumption, this method has been recognized as a good way to reduce the bias (Bound and Solon 1999). Another critique of the use of the twin-estimator is its sensitivity to measurement errors (Griliches 1979; Ashenfelter and Krueger 1994; Neumark 1994; Bound and Solon 1999; Light and Flores-Lagunes 2006). From the point of view of external validity, we may wonder whether twins are representative of the whole population: we know, for example, that they have lower weight at birth, they are more

likely to have problems in language, and they are brought up differently than other children (Mowrer 1954; Mittler 1971; Stewart 2000; Schieve et al. 2002).

When there are not enough twins (this is always the case in survey data), researchers have used siblings for the same purpose: trying to separate the effect of interest from the unobservable characteristics of the family (Behrman and Wolfe 1989; Neumark and Korenman 1994; Altonji and Dunn 1996; Ashenfelter and Zimmerman 1997; Aaronson 1998; Ermisch and Francesconi 2000). Using siblings instead of twins has some advantages but also evident shortcomings. On one hand, they all have the same family background, experience similar lifetime events (but at different times), are more representative of the population, and can provide larger samples and more precise estimates. On the other hand, they do not fully share the same genes and the results obtained are then potentially biased.

Another strategy used to eliminate, or at least reduce, the bias in the estimation of the intergenerational transmission of education is to use a sample of adopted children. In fact, there is no genetic transmission of ability between adopted children and their adoptive parents. In this case, a relationship between parental and children's education should reveal a causal link between the two, while the comparison with estimates obtained from own-born children can be suggestive of the importance of the family unobservables. The most common criticism of this strategy is that children are not randomly assigned: if adoption authorities have information on children's natural parents, they can use it to match children to adoptive couples. Another criticism derives from the fact that adoptive parents are not representative: they are, on average, older, better educated, and more motivated than the overall population. This could threaten the generalization of the results.

Finally, other studies have used a reform of the educational system as an instrument for parental education. For example, increasing the age of compulsory schooling lengthens the years of education exogenously. Exploiting this exogenous variation, we can observe whether the children of parents whose education has increased due to the reform achieve more in the education system than children of parents not influenced by the reform. There is no risk of bias in this case, but results cannot be generalized to the whole population if the reform of the schooling system involves only one part of the educational distribution.

There are a number of papers which use different strategies, data, outcomes, and control variables. I focus on those with "years of schooling" as the outcome variable and which take into account the level of education of both parents simultaneously. These studies are most comparable with the work carried out in this paper. Since the aim is to explain variation in the statistical significance of results, I indicate that a parameter estimate is significant at 10% level with one asterisk, at 5% with two asterisks, and at 1% with three asterisks.

A seminal work making use of twins in the study of the intergenerational transmission of education was published by Behrman and Rosenzweig in 2002. They use a sample of monozygotic twins drawn from the Minnesota Twin Register. Information was obtained through a mail survey. The data contain

424 twin-mothers and 244 twin-fathers. They find a positive effect of father's schooling on children's years of schooling (0.340**) and a negative effect of mother's education (-0.263^*). They suggest that this pattern is consistent with the fact that women's time is an important determinant of children's outcomes: the potential positive effect of mother's education is offset by the fact that more educated women spend more time in the labor market and less time with their children. Antonovics and Goldberger (2005) cast doubt on the construction of the dataset used by Behrman and Rosenzweig.¹ After what they consider to be an appropriate cleaning of the dataset, they have a sample of 90 pairs of twin-mothers and 47 pairs of twin-fathers. But the striking difference between genders remains: the effect of father's education is positive (0.477—standard error not available) and the effect of mother's education is close to zero (-0.003 —standard error not available).

To estimate the effect of mother's schooling on children's schooling, Plug (2004) controls for the effect of unobserved inherited abilities, obtaining identification by using adopted children instead of twins. He uses information collected in 1992 from 610 students who graduated from high schools in Wisconsin in 1957 and also finds a strong and positive effect of father's education (0.209***) and little effect of mother's education (0.089). He proposes different explanations for this result: some are substantive (better educated women spend less time with their children, differences in upbringing between own children and adopted children, adopted children different from other children in ways related to maternal schooling effects); some are more related to the design of the analysis (measurement error, heterogeneity with respect to the age adopted children enter their adoptive families, selection of highly educated mothers and consequently little variance in their education). Using Swedish adoption data, Björklund et al. (2006) are not only able to remove the family unobservable characteristics from the effect of interest but also able to distinguish between prebirth factors (genetics and prenatal environment) and postbirth environmental factors. They exploit the fact that Swedish register data contain information for both biological and adoptive parents. They have information on both pairs of parents for 2,125 adopted children. The effect of adoptive father's education is positive (0.094***) while that of adoptive mother's is small and insignificant (0.021).

Black et al. (2005) use a reform of compulsory schooling in Norway in the years 1961–1972 as an instrument for parental education. This reform resulted in 2 years more of schooling. They use administrative data linked with the municipalities which implemented the reform, year by year. They find a significant effect of parental schooling only when selecting low educated parents (the ones

¹Antonovics and Goldberger (2005) criticize that the sample used by Behrman and Rosenzweig (2002) includes very young children (also younger than 18 years old) for whom the final level of education cannot be observed and has been substituted with parents' expectations. However, De Hann and Plug (2009) propose and compare different methodologies to treat censored observations of the education variable and conclude that using parental expectations if they were realizations seems to deal relatively well with censoring problems.

most likely influenced by the reform): a positive effect of mother's education (0.122***) but a negligible effect of father's education (0.041).

These conflicting results obtained using various identification strategies and different datasets raise the question of what drives the differences. Are they data specific or do they depend on identification strategy? Holmlund et al. (2008) use different identification strategies with the same source of data, reaching the conclusion that it is the source of identification that matters. They select from Swedish register data parents born between 1943 and 1955, whose experience of the reform of compulsory schooling depended on the municipality of residence. They find a positive result of mother's education (0.150**), while they do not find any significant corresponding effect for father's education (0.019). Using information from around 4,000 foreign-born children adopted by parents of the 1943–1955 birth cohorts, they find small but positive effects of both father's (0.026**) and mother's education (0.034***). Finally, they have 5,586 children of twin-mothers and 4,061 of twin-fathers (only half of them monozygotic), from which they estimate a mother's effect equal to 0.038 and a father's effect equal to 0.110***.

This paper, by using rich data on Norwegian twins, aims at adding empirical evidence in this field of research by exploring how sensitive are results to the sample size and to the part of the population selected for the analysis.

3 Norwegian parents and children

The first step of my empirical research is to replicate the analyses carried out in previous work (Behrman and Rosenzweig 2002; Antonovics and Goldberger 2005; Holmlund et al. 2008) in order to provide comparable results. The twin-estimator, which is used to estimate the intergenerational schooling effect, indicates whether the child of the more educated twin obtains more schooling qualifications, controlling for the unobservable characteristics transmitted by the parent.

The informational basis for the empirical analysis is a register household panel dataset covering the entire resident population of Norway for the years 1993–2001. The dataset contains information on household size and composition as well as individual information such as place of residence, date of birth, educational level, and work status.

In this paper, twins are defined as people of the same sex, born in the same calendar year and month² from the same parents. In order to be included in the sample, both twins must have at least one child aged over 22 in 2001.³

²Register data provide a personal identifier from which we can extract month and year of birth but not the day. Therefore, it is possible that one child could be born just before the stroke of midnight on the last day of the month and his/her own twin just after midnight on the next day and the next month. However, we are probably missing very few and random cases.

³Holmlund et al. (2008) selected children aged over 22, Antonovics and Goldberger (2005) from 18 years old, and Behrman and Rosenzweig (2002) include also children younger than 18.

The data do not allow us to distinguish between monozygotic and dizygotic twins. Monozygotic twins have exactly the same genetic code, while dizygotic twins share the same proportion of the genes just like other siblings and differ from other siblings only by being born on the same day. Without distinguishing them, but only selecting same-sex pairs, we know that half of them are probably identical monozygotic twins.⁴

The twin-estimator is criticized for assuming random education between twins, for being sensitive to measurement error, and because twins may be not representative of the whole population. This paper does not debate the assumption of randomness of twins' education, but considers the twin-estimator at least as a valid method to reduce the bias. To investigate whether twins may be considered representative of the population, the paper provides some descriptive statistics comparing twins and the general population. Because of the use of the registers, measurement error should be minimal with these data.

The sample of twin-mothers is made of 1,609 mothers in 804 families who have 3,009 children aged over 22 in 2001; the sample of twin-fathers is composed of 1,606 fathers in 802 families with 3,086 children over 22.

In Table 1, I compare twins' families with nontwins' families, looking at their background and at their children's outcomes in terms of education, work, welfare dependence, and family composition. I select parents born in the same cohorts as twins and with at least one child more than 22 years old in 2001. For both parents and children, we have information on schooling, earnings, social transfers, and self-employment, which are measured in 1993 for parents and in 2001 for children over 22. The levels of education are transformed into years of schooling, according to the maximum level of education obtained.⁵ At the twin's/parent's level, we can obtain the number of siblings they had in their parental household and the number of children they have had during their life.⁶ Since women may have children from different men, and vice versa, information about the other parent's schooling and divorce is measured at the child's level.

Descriptive statistics are reported in Table 1 separately for mothers and fathers, twin-parents, and parents from the overall population. Comparing twins and general parents, we mainly observe age differences: parents from the overall population are older and, consequently, with older children than twin-parents. Twin-fathers earn more and receive a smaller amount of transfers, which may result from their higher likelihood to be still in the labor market, given their younger age. But there are no differences at all in schooling and income sources of their children. Therefore, we may be fairly confident that

⁴Also, Holmlund et al. (2008) cannot distinguish between monozygotic and dizygotic twins, while the sample employed by Behrman and Rosenzweig (2002) is composed only of monozygotic twins.

⁵Source: Norwegian Standard Classification of Education, Revised (2000), Official Statistics of Norway, C 751, Statistics Norway.

⁶The number of siblings and children is drawn from the register data in the year 1993, which is the first wave available. Therefore, both siblings and children need to be alive in 1993 to be observed and included in the samples.

Table 1 Descriptive statistics on Norwegian parents and children

	Mothers		Fathers	
	Twins	Population	Twins	Population
Parent's level				
Age (1993)	44.8 (6.8)	47.1 (8.6)	47.9 (7.5)	50.8 (9.4)
Number of siblings (1993)	3.66 (3.88)	3.72 (4.99)	3.53 (4.02)	3.89 (5.37)
Years of schooling (1993)	10.9 (2.1)	11.2 (2.2)	11.5 (2.6)	11.6 (2.6)
Earnings (€, 1993)	13,207 (10,325)	12,382 (10,312)	22,888 (17,826)	20,423 (19,381)
Transfers (€, 1993)	3,091 (4,321)	3,210 (4,275)	2,416 (5,084)	3,437 (6,025)
Self-employed (1993)	0.102	0.097	0.223	0.260
Number of children (1993)	2.45 (0.95)	2.42 (1.02)	2.44 (0.95)	2.51 (1.07)
Parents, observations	1,609	278,390	1,606	303,703
Child's level				
Age (2001)	27.3 (7.3)	29.4 (8.8)	28.0 (7.5)	29.7 (9.3)
Years of schooling (2001)	12.9 (2.4)	12.9 (2.4)	12.9 (2.4)	12.9 (2.4)
Other parent's schooling (1993)	11.6 (2.6)	11.5 (2.6)	11.0 (2.2)	11.0 (2.2)
Divorce (1993)	0.200	0.176	0.185	0.159
Earnings (€, 2001)	25,241 (18,104)	25,540 (19,289)	25,513 (17,611)	25,740 (19,360)
Transfers (€, 2001)	3,239 (5,524)	3,393 (5,673)	3,178 (5,349)	3,365 (5,655)
Self-employed (2001)	0.079	0.097	0.087	0.105
Children, observations	3,945	674,507	3,918	764,256
Children over 22, observations	3,009	545,523	3,086	618,550

Descriptive statistics on Norwegian parents and children are shown separately for twins and the overall population. Average values with standard deviations in parentheses. “Self-employed” is a dummy variable indicating whether part of the income is from self-employment work. “Number of children” is defined for children of any age. “Age” at the child’s level is measured for all children while the other variables at the child’s level are only summarized for children over 22

results obtained by twins, in this specific field, may be generalized to the overall population.

4 The use of twins for estimating the intergenerational transmission of schooling

I define Y as the years of schooling of the twin’s child, X as the years of schooling of the twin-parent, and Z as other factors which may influence the child’s education. For each child i in the family j , we have:

$$Y_{ji} = \beta X_{ji} + Z'_{ji} \alpha + u_j + \varepsilon_{ji} \quad (1)$$

where β is the effect of parental education on the child's years of schooling, α the effect of other factors, u are the unobservable characteristics shared in the twins' family j , and ε is the error term assumed to be white noise. A pooled regression of Y on X and Z (cross-section) is not appropriate since it ignores the unobservable characteristics u shared in each twins' family, which may be correlated with parental education. We can eliminate u from the equation, differencing the data in the following way:

$$(y_{ji} - \bar{y}_j) = \beta (x_{ji} - \bar{x}_j) + (z_{ji} - \bar{z}_j)' \alpha + (\varepsilon_{ji} - \bar{\varepsilon}_j) \quad (2)$$

where \bar{y}_j is the average years of schooling of twins' children in family j , \bar{x}_j is the average years of schooling of twin-parents in family j , and \bar{z}_j are other average characteristics of the twins' family j .

Among the factors Z which may affect children's schooling, I control for the other parent's characteristics, parental divorce, gender, and age of the child. The aim of considering the other parent's characteristics is to control for assortative mating. Following Behrman and Rosenzweig (2002), I include, beyond the other parent's years of schooling, the earnings endowment which is the part of his/her earnings that does not depend on schooling, i.e., the residual from the earnings equation.⁷ The dummy variable "divorce" is equal to 1 when the two parents are not living together in 1993. Finally, the inclusion of the variable "age" helps to control for performance of younger students, less likely to have reached the highest levels of education.

In Table 2, results from the twin-estimator are presented. In the ordinary least squares (OLS) estimation, all variables with the exception of age are highly significant and in the expected direction. In the twin estimations, the effect of twin's schooling is reduced, particularly for mothers, but is still positive and significant for both mothers (0.096**) and fathers (0.158***). For the first time, applying this identification strategy, a positive effect of mother's education has been found. Moreover, the effect of mother's schooling is not found to be significantly different from the effect of father's education.

However, one limitation of the data is that we cannot distinguish between monozygotic and dizygotic twins. To the extent to which the unobservables which influence the child's schooling attainment are written in the genes, this lack of information may bias the estimate, since dizygotic twins do not share all the same genetic code. In order to understand how much this matters, I select siblings who were born very close together, estimate the intergenerational transmission of schooling, and compare the estimate with the one obtained using twins, who may be considered a mixed group of identical twins and

⁷The omission of this variable may bias the effect of the twins' schooling, if earnings of the other parent are correlated with twin's education (e.g., richer men get married with better educated women). On the other hand, the inclusion of observed earnings may bias downwards the effect of the other parent's schooling, since observed earnings depend on education. Register data contain annual earnings but not hours of work; therefore, a yearly earnings equation has been estimated and residuals included in the model.

Table 2 The intergenerational schooling effect (OLS and twin-estimator)

	Mothers		Fathers	
	OLS	Twin-estimator	OLS	Twin-estimator
Twin's schooling	0.242*** (0.023)	0.096** (0.041)	0.214*** (0.018)	0.158*** (0.033)
Other parent's schooling	0.216*** (0.018)	0.124*** (0.025)	0.215*** (0.021)	0.161*** (0.030)
Other parent's earnings endowment	0.009*** (0.002)	0.009*** (0.003)	0.023*** (0.005)	0.017*** (0.007)
Divorce	-0.529*** (0.100)	-0.361*** (0.152)	-0.748*** (0.101)	-0.260* (0.141)
Age	0.008 (0.007)	0.018 (0.013)	0.002 (0.006)	-0.010 (0.012)
Female	0.504*** (0.079)	0.502*** (0.085)	0.447*** (0.078)	0.466*** (0.083)
Constant	7.433*** (0.341)	9.755*** (0.643)	7.989*** (0.344)	9.486*** (0.601)
Children, observations	3,009	3,009	3,086	3,086
Families, observations	804	804	802	802

Coefficients are shown with standard errors in parentheses

***significant at the 1% level,

**significant at the 5% level,

*significant at the 10% level

siblings born on the same day. I select siblings whose difference in age is between 9 and 13 months. This allows us to have about the same sample size. Results are reported in Table 3: the effect of mother's schooling is positive (0.139***) and not significantly different from the one obtained from twins; the effect of father's education is positive and smaller (0.123***) but not significantly different from the one obtained by twins. By knowing that half of twins are monozygotic and half dizygotic and assuming that the coefficients are normally distributed, the maternal effect for identical twins should be equal to 0.052 (SE = 0.089) while the paternal effect should be 0.192 (SE = 0.072).⁸

There is no longer enough evidence to say that mother's education has an impact on children's education. Nevertheless, the mother's and father's effects have not yet proved to be different. Obviously, these results suffer from imprecision and by themselves reinforce the motivation of the paper. How much are the estimates of the intergenerational schooling coefficient robust to sampling issues? Are estimates influenced by small sample size? Should we prefer siblings whose estimates are potentially biased but more precise? Are the effects heterogeneous along the distribution of parents' education?

In order to assess the effect of sample size on the estimates of intergenerational transmission of schooling, I work through simulations: I assume

⁸The average effect of the twins-parents' education for the mixture of monozygotic and dizygotic twins is $\beta_{MIX} = 0.5\beta_M + 0.5\beta_D$. In this exercise, β_D is approximated by the effect estimated with the sample of close siblings.

Table 3 The intergenerational schooling effect (sibling-estimator)

	Siblings (9–13 months of age difference)	
	Mothers	Fathers
Sibling's schooling	0.139*** (0.036)	0.123*** (0.030)
Other parent's schooling	0.134*** (0.026)	0.187*** (0.031)
Other parent's earnings endowment	0.006** (0.003)	0.020*** (0.006)
Divorce	-0.602*** (0.130)	-0.392*** (0.134)
Age	-0.014 (0.014)	-0.012 (0.012)
Female	0.361*** (0.086)	0.300*** (0.085)
Constant	10.072*** (0.637)	9.615*** (0.582)
Children, observations	2,957	3,052
Families, observations	823	799

The samples are composed of pairs of siblings whose distance in age is between 9 and 13 months. Coefficients are shown with standard errors in parentheses

***significant at the 1% level,

**significant at the 5% level,

*significant at the 10% level

the sample of Norwegian twin-parents to be the population of reference from which I draw samples of different size, estimate the effect of parental schooling, and count how many times I would reject the null hypothesis of zero effect.⁹ The results are summarized in Fig. 1. The probability of rejecting the hypothesis of a null effect is increasing in the number of families in the sample, as expected, but the difference between genders is remarkable: given relatively small samples and effects rather close to zero, typical of this kind of study, we are always more likely to reject the null hypothesis at the 1% level for fathers than at the 10% level for mothers. It is also hard to say about the difference between mother's and father's effects with these sample and effect sizes: by doubling the number of families (around 1,600 instead of 800), the probability of rejecting the hypothesis that the maternal effect is equal to the paternal effect is only 0.222 at the 5% level. This paper confirms the important role of father's education, found in many twins' study, but also supports the idea that results suffer from little precision, especially in the mother's case, since the mother's effect is closer to zero (but not equal to zero in this paper) than father's. Behrman and Rosenzweig (2002), followed by Antonovics and Goldberger (2005), use very small samples which could explain the zero effect of mother's education. On the other hand, Holmlund et al. (2008), using twins but a larger sample size, did not find any effect.

⁹I draw 1,000 samples of each different size (from 100, 200, . . . , to 1,600 families).

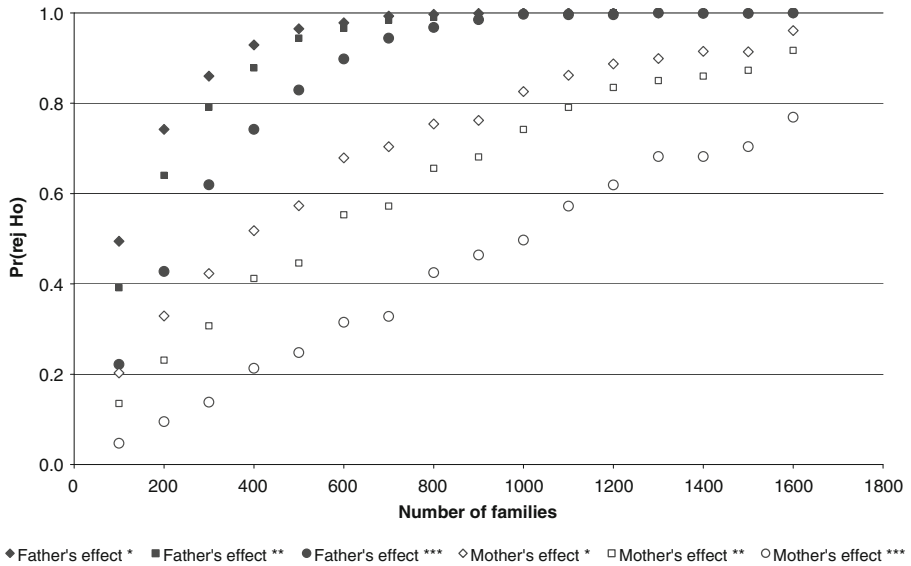


Fig. 1 The impact of sample size on the estimated intergenerational schooling effect. The probability of rejecting the hypothesis that the intergenerational schooling effect is equal to zero (***)significant at the 1% level, **significant at the 5% level, *significant at the 10% level) is shown in the figure, given that the true effect is equal to 0.096 for mothers and 0.158 for fathers, obtained by using the twin-estimator and controlling for other parent's schooling and earnings endowment, divorce, child's age, and gender

In order to study heterogeneous effects of an increase in schooling along the distribution of parents' education, I divide the sample of twins into two parts: one part where both twins are lower educated (primary or lower secondary school) and another part where both are higher educated (upper secondary or tertiary schooling). In Table 4, we observe a strong and significant effect of father's education, but only in the top part of the distribution. On the other hand, no effect of mother's education is found in the top part, while there is indication of a positive but still insignificant effect of an increase in schooling for lower educated mothers.

Using the same simulation setting,¹⁰ I measure how likely we are to reject the null hypothesis by varying the number of families in the sample (Fig. 2) and by focusing separately on the top and bottom part of the distribution of parents' education. Differences are huge. When looking at highly educated parents, we are very likely to find a significant effect of father's education even with a small sample size, while the probability of observing a significant effect of mother's education is very small, independently of the number of observations in the analysis. When selecting lower educated parents, the

¹⁰I draw 1,000 samples of each different size (from 100, 200, ..., to 1,600 families), from which I select families where both twins are low/high educated.

Table 4 Heterogeneous intergenerational schooling effects (twin-estimator)

	Twin-mothers		Twin-fathers	
	Low educated	High educated	Low educated	High educated
Twin's schooling	0.125 (0.084)	0.100 (0.119)	0.073 (0.098)	0.291*** (0.078)
Other parent's schooling	0.135*** (0.031)	0.054 (0.076)	0.207*** (0.048)	0.178*** (0.056)
Other parent's earnings endowment	0.012*** (0.004)	0.001 (0.008)	0.022** (0.009)	0.014 (0.017)
Divorce	-0.438** (0.176)	0.634 (0.512)	-0.285 (0.198)	-0.682** (0.341)
Age	-0.009 (0.014)	0.269*** (0.060)	-0.038** (0.015)	0.060* (0.033)
Female	0.417*** (0.097)	0.855*** (0.298)	0.353*** (0.112)	0.772*** (0.203)
Constant	10.057*** (1.022)	4.768** (2.662)	10.667*** (1.150)	5.469*** (1.573)
Children, observations	2,282	270	1,588	609
Families	590	79	400	174

Samples are composed of twins who are both low/high educated. Coefficients are shown with standard errors in parentheses

***significant at the 1% level,

**significant at the 5% level,

*significant at the 10% level

probability of rejecting the hypothesis that the effect is zero is very small for fathers, while larger and increased by a larger sample size for mothers.

How can we interpret these results? The results in this paper seem to confirm those obtained using different strategies than twins. Studies which make use of adoption are likely to use small samples of parents better educated than average. They find a positive effect of father's schooling and no effect or a small effect of mother's schooling (Plug 2004; Björklund et al. 2006), which are consistent with the effects found in this paper for highly educated parents. On the other hand, studies which make use of compulsory schooling reform as an instrument (Black et al. 2005; Holmlund et al. 2008) only find a positive effect of mother's education. This paper does not find a significant effect of schooling for lower educated mothers but simulations on the power of the test show that the probability of rejecting the null hypothesis increases considerably with the number of families in the sample. And, opposite to the studies which employ twins and adopted children, studies which use a schooling reform as an instrument have quite large samples.¹¹

Both sample size and heterogeneous effects along the distribution of parents' education are important for explaining different results by gender and also help to reconcile the apparently conflicting results in the literature when using different identification strategies.

¹¹Black et al. (2005) have a sample of low educated mothers of around 40,000 cases.

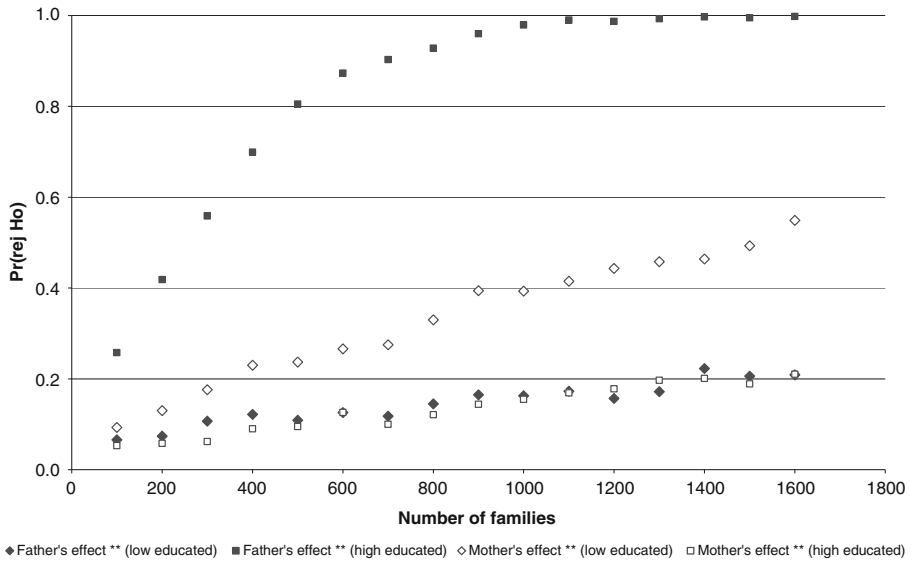


Fig. 2 The impact of sample size and selection on the estimated intergenerational schooling effect. The probability of rejecting the hypothesis that the intergenerational schooling effect is equal to zero (**significant at the 5% level) is shown in the figure, given that the true effect is equal to 0.125 for low educated mothers, 0.100 for high educated mothers, 0.073 for low educated fathers, and 0.291 for high educated fathers, obtained by using the twin-estimator and controlling for other parent's schooling and earnings endowment, divorce, child's age, and gender; total samples are composed of the number of families indicated on the X -axis, before selecting twins who are both low/high educated (on average, 73% of female twins are both lower educated and 10% higher educated; 50% of male twins are both lower educated and 22% higher educated)

5 The use of siblings for understanding the intergenerational transmission of schooling

The sibling-estimator is used with the same aims as the twin-estimator when the sample of twins is not large enough. Siblings share most of the characteristics of the family background but not fully the genes and, depending on the difference in age, experience events at different times. One aim of this section is to show whether we would arrive at the same conclusions by using siblings instead of twins, and in this way, provide a benchmark for other research. However, the study of siblings may also be informative by itself: the size of the bias we may get from pairs of siblings is indicative of the source of what we have generally called family unobservables. If it is ability or other characteristics in the genetic code, which creates correlation between parents' and children's education, then we should find a substantial difference between twins' estimates and close siblings' estimates, since close siblings experience events almost at the same time but differ in part of their genes. Instead, if it is the time of the events which matters, we expect to find that the difference between twins' and close siblings' estimates is smaller than the difference between twins' estimates and estimates based on siblings born further apart.

In making this comparison, we need to recall that samples in this paper are composed of both monozygotic and dizygotic twins.

Table 5 shows the estimates of the intergenerational transmission of schooling using siblings. The samples are composed of pairs of siblings with different distances in age: the closest were born not more than 13 months apart, the furthest between 4 and 5 years. We observe a positive and significant effect of mother's and father's education on children's education with all samples of siblings. Moreover, there are no significant differences between the mother's and father's effect at any level of distance in age. We would have come to the same conclusions: not only do the estimates show the same direction and significance, but sibling estimates are never significantly different from twin estimates. By comparing twins, closest siblings, and furthest siblings, there is no strong evidence of bias by ability or by the time at which events happen.

However, results become puzzling when the analysis is repeated by focusing separately on the bottom and top parts of the distribution of parents' education (Table 6). Mother's education is confirmed to matter in the bottom part of the distribution. Furthermore, sister estimates never significantly differ from twin estimates. The effect of father's education is still strong in the top part of the distribution, but only when we consider brothers born no more than 13 months apart. One additional year of schooling for highly educated brothers has a significantly smaller effect than for highly educated twins. The effect becomes smaller as long as the distance in age between siblings increases. Why

Table 5 The intergenerational schooling effect (twin- and sibling-estimators)

	Mothers	Fathers
Twins	0.096** (0.041) [804]	0.158*** (0.033) [802]
Siblings (9–13)	0.139*** (0.036) [823]	0.123*** (0.030) [799]
Siblings (14–24)	0.128*** (0.012) [7,783]	0.122*** (0.009) [8,588]
Siblings (25–36)	0.123*** (0.011) [7,922]	0.140*** (0.009) [8,792]
Siblings (37–48)	0.126*** (0.012) [6,962]	0.132*** (0.009) [7,904]
Siblings (49–60)	0.125*** (0.013) [5,706]	0.137*** (0.011) [6,553]

Samples of siblings are composed of pairs of siblings whose distance in age is between 9 and 13 months, 14 and 24, ...; other parent's schooling and earnings endowment, divorce, child's age and gender are included in the regressions but coefficients are not reported. Coefficients are shown with standard errors in parentheses (***)significant at the 1% level, **significant at the 5% level, *significant at the 10% level) and number of families in square brackets

Table 6 Heterogeneous intergenerational schooling effects (twin- and sibling-estimators)

	Mothers		Fathers	
	Low educated	High educated	Low educated	High educated
Twins	0.125 (0.084) [590]	0.100 (0.119) [79]	0.073 (0.098) [400]	0.291*** (0.078) [174]
Siblings (9–13)	0.206*** (0.075) [589]	−0.036 (0.122) [67]	0.116 (0.086) [386]	0.168** (0.078) [139]
Siblings (14–24)	0.199*** (0.024) [5,490]	0.002 (0.039) [708]	0.164*** (0.026) [4,238]	0.112***{***} (0.023) [1,591]
Siblings (25–36)	0.163*** (0.024) [5,484]	0.070* (0.036) [833]	0.188*** (0.025) [4,285]	0.120***{***} (0.022) [1,655]
Siblings (36–48)	0.118*** (0.026) [4,825]	0.105** (0.043) [678]	0.181*** (0.027) [3,738]	0.101***{***} (0.022) [1,507]
Siblings (49–60)	0.176*** (0.028) [4,044]	0.110** (0.046) [499]	0.195*** (0.030) [3,289]	0.099***{***} (0.025) [1,167]

Samples are composed of twins or siblings who are both low/high educated; samples of siblings are composed of pairs of siblings whose distance in age is between 9 and 13 months, 14 and 24, ...; other parent's schooling and earnings endowment, divorce, child's age and gender are included in the regressions but coefficients are not reported. Coefficients are shown with standard errors in parentheses (***)significant at the 1% level, **significant at the 5% level, *significant at the 10% level) and number of families in square brackets; significant differences between estimates from siblings and twins are reported in curly brackets (***)significant at the 1% level, **significant at the 5% level, *significant at the 10% level)

is that? One possible explanation, which we may find in the literature, is that siblings' parents may have compensated for differences in children's abilities by investing more in the less able child (Ermisch 2003), creating a negative correlation between ability and education, which biases downwards the effect of parental education.

6 Conclusions

In this paper, monozygotic and dizygotic twins have been used to estimate the intergenerational transmission of schooling. The paper confirms the strong effect of father's education found in previous twins' studies, but also finds a positive and significant effect, though smaller, of mother's education.

The contributions of the paper to this field of research are three. First, it assesses the impact of small sample size on the robustness of the results. Given the size of the intergenerational effects found in this paper, we would need at least 1,000 pairs of twins to be confident that an insignificant coefficient is only due to a true zero effect.

Second, by focusing on the effect of one more year of education separately for children of lower and higher educated parents, the paper narrows the distance between the results provided by the three different identification strategies. When I employ pairs of highly educated parents, I only find a positive effect of father's education, which is what was often found in previous papers when using samples of parents of adopted children who are, on average, better educated than the overall population. On the other hand, when I employ pairs of lower educated parents, I only observe a potentially positive impact of mother's education, which confirms results from papers which exploit a reform of compulsory schooling as an exogenous variation.

Third, the paper compares, in a systematic way, the estimates of the intergenerational transmission of schooling effect obtained from samples of twins and samples of siblings. If there were no differences, we could use siblings from survey data and make use of interesting information on parenting style, parental time spent with the child, and other family characteristics. However, results provided by the sibling-estimator tend to underestimate the positive effect of an additional year of education of highly educated fathers. One plausible explanation is that siblings' parents may have compensated for differences in children's abilities by investing more in the less able child (Ermisch 2003). This would create a negative correlation between ability and education, biasing downwards the effect of parental education, since ability between siblings is not taken into account as well as it is between twins. This result is in contrast with that found by Behrman et al. (1994): they observe a reinforcement behavior in the family, by using samples of distinguishable monozygotic and dizygotic twins. However, the compensation behavior hypothesized in this paper only concerns parents of high educated brothers and we would not reach the same conclusion by looking at lower educated brothers, while support the idea of a reinforcement if any. Moreover, by using siblings instead of dizygotic twins, parents have more time to learn about the differences in abilities of their children and may feel less unfair in treating siblings in a different way.

This is left for further research. Whether or not the compensation/reinforcement behavior is the one at work, empirical results indicate that samples of twins should be preferred to samples of siblings in the study of intergenerational transmission of schooling.

Acknowledgements I thank my Ph.D. supervisor John Ermisch, Rolf Aaberge, Richard Berthoud, Ugo Colombino, Francesco Figari, Helena Holmlund, Cheti Nicoletti, and Steve Pudney for their comments. I am very grateful to two anonymous referees whose precious comments and suggestions improved my work and to Deborah Cobb-Clark for the invaluable advice. The financial support received by the Norwegian Research Council and the European Research Council are gratefully acknowledged. Any error should be attributed to the author.

References

- Aaronson D (1998) Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *J Hum Resour* 33(4):915–946

- Altonji JG, Dunn TA (1996) Using siblings to estimate the effect of school quality on wages. *Rev Econ Stat* 78(4):665–671
- Antonovics KL, Goldberger AS (2005) Does increasing women's schooling raise the schooling of the next generation? *Comment Am Econ Rev* 95(5):1738–1744
- Ashenfelter O, Krueger A (1994) Estimates of the economic return to schooling from a new sample of twins. *Am Econ Rev* 84(5):1157–1173
- Ashenfelter O, Zimmerman DJ (1997) Estimates of the returns to schooling from sibling data: fathers, sons, and brothers. *Rev Econ Stat* 79(1):1–9
- Behrman JR, Rosenzweig MR (2002) Does increasing women's schooling raise the schooling of the next generation? *Am Econ Rev* 92(1):323–334
- Behrman JR, Wolfe BL (1989) Does more schooling make women better nourished and healthier? Adult sibling random and fixed effects for Nicaragua. *J Hum Resour* 24(4):644–663
- Behrman JR, Rosenzweig MR, Taubman P (1994) Endowments and the allocation of schooling in the family and in the marriage market. *J Polit Econ* 102(6):1131–1174
- Björklund A, Lindahl M, Plug E (2006) The origins of intergenerational associations: lessons from Swedish data. *Q J Econ* 121(3):999–1028
- Black SE, Devereux PJ, Salvanes KG (2005) Why the apple doesn't fall far: understanding intergenerational transmission of human capital. *Am Econ Rev* 95(1):437–449
- Bound J, Solon G (1999) Double trouble: on the value of twin-based estimation of the return to schooling. *Econ Educ Rev* 18(2):169–182
- De Hann M, Plug E (2009) Estimating intergenerational schooling mobility on censored samples: consequences and remedies. *J Appl Econometrics* (forthcoming)
- Ermisch J (2003) *An economic analysis of the family*. Princeton University Press, Princeton
- Ermisch J, Francesconi M (2000) Educational choice, families, and young people's earnings. *J Hum Resour* 35(1):143–176
- Griliches Z (1979) Siblings models and data in economics: beginning of a survey. *J Polit Econ* 87(5):S37–S64
- Haveman R, Wolfe B (1995) The determinants of children's attainments. *J Econ Lit* 33(4):1829–1878
- Hertz T, Jayasundera T, Piraino P, Selcuk S, Smith N, Verashchagina A (2007) The inheritance of educational inequality: international comparisons and fifty-year trend. *BE J Econ Anal Policy* 7(2):article 10
- Holmlund H, Lindahl M, Plug E (2008) Estimating intergenerational schooling effect: a comparison of methods. *IZA Discussion Paper* 3630
- Light A, Flores-Lagunes A (2006) Measurement error in schooling: evidence from samples of siblings and identical twins. *BE J Econ Anal Policy* 5(1):article 14
- Mittler P (1971) *The study of twins*. Penguin Books, Baltimore
- Mowrer ER (1954) Some factors in the affectional adjustment of twins. *Am Sociol Rev* 19(4):468–471
- Neumark D (1994) Biases in twin estimates of the return to schooling: a note on recent research. *National Bureau of Economic Research Technical Working Paper* 158
- Neumark D, Korenman S (1994) Sources of bias in women's wage equations: results using sibling data. *J Hum Resour* 29(2):379–405
- Plug E (2004) Estimating the effect of mother's schooling on children's education using a sample of adoptees. *Am Econ Rev* 94(1):358–368
- Schieve LA, Meikle SF, Ferre C, Peterson HB, Jen G (2002) Low and very low birth weight in infants conceived with use of assisted reproductive technology. *N Engl J Med* 346(10):731–737
- Stewart EA (2000) *Exploring twins: towards a social analysis of twinship*. McMillan, London