

ESTIMATING INTERGENERATIONAL SCHOOLING MOBILITY ON CENSORED SAMPLES: CONSEQUENCES AND REMEDIES

MONIQUE DE HAAN^a AND ERIK PLUG^{a,b,*}

^a *Amsterdam School of Economics, University of Amsterdam, Amsterdam, Netherlands*

^b *Tinbergen Institute, Amsterdam, Netherlands; IZA, Bonn, Germany*

SUMMARY

In this paper we estimate the impact of parental schooling on child schooling, focus on the problem that children who are still in school constitute censored observations, and evaluate three solutions to it: replacement of observed with expected years of schooling, maximum likelihood approach, and elimination of all school-aged children. Using intergenerational data from the Wisconsin Longitudinal Study we test how the three correction methods deal with censored observations. The one that treats parental expectations as if they were realizations seems to fix the censoring problem quite well. Copyright © 2009 John Wiley & Sons, Ltd.

Received 7 December 2007; Revised 20 March 2009

1. INTRODUCTION

Most empirical studies on intergenerational mobility estimate a version of the following model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t \quad (1)$$

where t is a generation index, Y_t and Y_{t-1} represent *realized* outcomes of child and parent, and u_t is a child-specific characteristic. In most studies the parameter of interest is β_1 , which measures the outcome association between parent and child. Estimating β_1 , however, puts strong requirements on data. In household surveys and censuses, in particular, the collection of information on realized outcomes of children is often problematic. Information on completed schooling, for example, is rarely available, for a number of reasons. Children who live with their parents are mostly still in school: their schooling information is by definition incomplete; and children who have completed their schooling have usually moved out of their parents' home: their schooling information is seldom collected.¹

In this paper we let Y be years of schooling, and focus on the problem that (some) children may still be in school at the time of data collection, which goes under the name of the censoring problem. We consider this a serious problem for three reasons. First, we cannot ignore censoring empirically because, if we do, least squares regression on censored samples would give us intergenerational persistency estimates that are too low. Second, we observe that censoring is a widely spread

* Correspondence to: Erik Plug, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. E-mail: e.j.s.plug@uva.nl

¹ Haider and Solon (2006) and Bölmarm and Lindquist (2006) consider the income analogue of our schooling example. Income information for children is difficult to collect because most children who live with their parents do not work; and even if income information is available, the intergenerational estimates will be biased downwards when the children's income is measured too early in life.

phenomenon. Of the recent studies that (aim to) estimate the causal effect of parents' schooling on children's schooling, almost all make use of samples with incomplete information on adult children. Among these studies are Behrman and Rosenzweig (2002, 2005); Plug (2004); Chevalier (2004); Black *et al.* (2005); Carneiro *et al.* (2007); Oreopoulos *et al.* (2006); and Maurin and McNally (2008). Third, the solutions offered to handle censored samples rely on assumptions that may not hold in practice, resulting in intergenerational mobility estimates that are biased.

Of course, the natural solution to the censoring problem is patience. If researchers were patient and could wait until all children in the censored sample finished their schooling to collect their data, we would not need to worry about censoring. Unfortunately, many researchers tend to be impatient. They are, presumably, more interested in the degree of intergenerational mobility among current generations than previous generations and are therefore willing to estimate parental schooling effects on censored samples using correction methods that do not always work.² Since the latter approach certainly merits serious consideration, it is important to know (more) about how applied correction methods deal with censored observations.

One procedure for testing how effective available correction methods are in treating censored schooling observations is to apply these censoring solutions to one particular dataset. This is what we do in this paper. We focus our attention on three correction methods that are currently in use: maximum likelihood approach, replacement of observed with expected years of schooling, and elimination of all school-aged children. We apply these three different methods to one particular dataset: the most recent version of the Wisconsin Longitudinal Study (henceforth often WLS).

The WLS collects information on a large group of students who graduated from Wisconsin high schools in 1957. In 1975, 1992 and 2004 the same students were contacted again and asked about their children's schooling. The questions cover three different school stages. In 1975 most children are in school: the sample includes information on expected schooling. In 1992 some of these children have not yet completed their schooling: the sample is a censored sample. In 2004 all children have left school: the sample contains information on completed schooling.

Our contributions are twofold. First, we present new estimates of the intergenerational mobility of schooling. With updated 2004 samples, we are able to estimate the ultimate mobility models in which censored observations are absent. Second, we examine the validity of the different solutions to deal with the problem of censored data. With the 1975 and 1992 samples, we estimate the impact of parental schooling on children's schooling, applying the various procedures to correct for censored observations and use the difference between ultimate and corrected mobility estimates as a validity indicator.

This paper continues as follows. Section 2 models the intergenerational mobility of schooling, focuses on the problem that children who are still in school generate censored observations, and provides some intuition on the various solutions to it. Section 3 provides a brief description of the WLS. Section 4 presents and compares the parameter estimates. Section 5 evaluates the correction methods and presents a number of robustness tests. Section 6 concludes.

2. MOBILITY MODELS USING CENSORED DATA

Much work on intergenerational schooling mobility has concentrated on estimating a version of the following model:

$$S_t = \beta_0 + \beta_1 S_{t-1} + \varepsilon_t \quad (2)$$

² The authors of the present paper plea guilty on being at least as impatient as anyone.

where t is a generation index, S_t and S_{t-1} represent the schooling of child and parent, usually measured as the number of years of completed schooling, and ε_t is a child-specific characteristic. The parameter β_1 measures the association between the schooling of parent and child. With information on S_t and S_{t-1} , the properties of the least-squares estimator are defined as

$$\text{plim } \beta_{1\text{LS}} = \text{cov}(S_t, S_{t-1})/\text{var}(S_{t-1}) = \beta_1 \quad (3)$$

A well-known problem in analyzing intergenerational schooling mobility is that information on the child's completed schooling is not always available. Some children are still in school at the time data are collected and create censored observations. To accommodate censored observations, we define a new variable S_t^c such that

$$\begin{aligned} S_t^c &= S_t \quad \text{if } d_t = 0, \\ S_t^c &\leq S_t \quad \text{if } d_t = 1 \end{aligned} \quad (4)$$

where S_t^c represents the child's years of schooling observed in the censored sample, and d_t denotes whether observations are censored ($d_t = 1$) or not ($d_t = 0$). If we would ignore censoring, and treat the children's observed years of schooling as their completed years, the estimation of S_t^c on S_{t-1} using ordinary least squares gives us a β_1 parameter that is too low. The intuition is as follows. We know that more schooled children (with more schooled parents) are more likely to be censored, and we know that observed school years are always smaller than or equal to the completed school years. Taken together, these observations imply that observed years of schooling covary less with parental years of schooling ($\text{cov}(S_t^c, S_{t-1}) \leq \text{cov}(S_t, S_{t-1})$). When we take a censored sample and apply least squares to estimate the child-parent equation

$$S_t^c = \beta_0^c + \beta_1^c S_{t-1} + \varepsilon_t^c \quad (5)$$

it follows naturally that the corresponding least squares estimator is biased toward zero, as

$$\text{plim } \beta_{1\text{LS}}^c = \text{cov}(S_t^c, S_{t-1})/\text{var}(S_{t-1}) \leq \text{cov}(S_t, S_{t-1})/\text{var}(S_{t-1}) = \beta_1 \quad (6)$$

Recent work on intergenerational mobility of schooling has taken various approaches to tackle the censoring problem: replacement of observed with expected years of schooling, maximum likelihood approach, elimination of all school-aged children, and intermediate school outcomes. Below we shortly discuss these different approaches.

2.1. Inserting Parental Expectations for Children Still in School

Behrman and Rosenzweig (2002) employ a mail survey, issued in 1994, to collect information on the families of identical twins born between 1936 and 1955, all drawn from the Minnesota Twin Registry (MTR). The survey contains information on the schooling of the twins, their parents and children, including information on expected schooling for children who had not yet completed their schooling; this is the case for more than 50% of their sample.³

³The American Economic Review provides data and programs for replication purposes online. From this source we have extracted the twin sample using data and programs of Antonovics and Goldberger (2005). We are able to trace 844 monozygotic twin parents with children. Of these 844 children, 428 were still in school in 1994.

Behrman and Rosenzweig replace their censored observations with parental expectations and treat these expectations as if they were school realizations for children with unfinished schooling. This gives the following school variable for the child:

$$\tilde{S}_t = \begin{cases} S_t & \text{if } d_t = 0, \\ S_t^e & \text{if } d_t = 1 \end{cases} \quad (7)$$

where S_t^e represents the school level the parent expects her child to complete. Suppose we model parental expectations about their children's completed years of schooling as follows:

$$S_t^e = S_t + \eta_t \quad (8)$$

where η_t is the error parents make in predicting their child's completed schooling.⁴ Combining (2), (7) and (8) leads to

$$\tilde{S}_t = \beta_0 + \beta_1 S_{t-1} + d_t \eta_t + \varepsilon_t \quad (9)$$

Applying least squares to the bivariate regression of \tilde{S}_t on S_{t-1} gives us the following probability limit of the slope coefficient:

$$\text{plim } \tilde{\beta}_{\text{ILS}} = \text{cov}(\tilde{S}_t, S_{t-1}) / \text{var}(S_{t-1}) = \beta_1 + \text{cov}(d_t \eta_t, S_{t-1}) / \text{var}(S_{t-1}) \quad (10)$$

Only if $\text{cov}(d_t \eta_t, S_{t-1})$ equals 0 does Behrman and Rosenzweig's original solution produce a consistent estimate of β_1 . If not, the validity of the method will depend on how much the prediction error correlates with parental education and on the number of censored observations. Whether or not $\text{cov}(d_t \eta_t, S_{t-1})$ equals 0 is an empirical issue, which we will put to the test later in this paper.

2.2. A Censored Regression Model

Plug (2004) exploits the 1992 wave of the WLS to estimate the effect of father's and mother's schooling on child's schooling using samples of biological and adopted children. In 1992 most parents are about 52 years old and many of their children have not yet finished their schooling (about 25% of the biological children and 40% of the adopted children). As we already mentioned, not taking censoring into account gives inconsistent estimates. Plug therefore uses a censored regression model, one of the standard procedures for handling censored observations. Assuming the conditional distribution of ε_t is normal with mean zero and constant variance σ^2 the likelihood function is

$$L(\theta) = \prod_{i=1}^N [\phi(S_i | S_{i-1}, \theta)]^{1-d_i} [1 - \Phi(S_i^c | S_{i-1}, \theta)]^{d_i} \quad (11)$$

where ϕ and Φ represent normal density and distribution functions, θ are the distribution parameters that include β_0 , β_1 and σ , and i indexes the family in which the child is born and raised. Maximization of (11) yields a consistent estimator of β_1 , unless the error distribution is incorrectly specified, being non-normally distributed or having a heteroskedastic variance.

⁴We omit subscript i here, but we do not assume that the prediction error is the same for all individuals, nor do we assume anything about the distribution of η_t .

2.3. Eliminating All School-Aged Children

Black *et al.* (2005) estimate the effect of parental schooling on child schooling using a reform in compulsory schooling in Norway during the 1960s and early 1970s to draw causal inferences. Because Black *et al.* focus on relatively young parents—only those between 42 and 53 years old are affected by the reform—many children have not yet finished their schooling by the time they appear in their sample. The authors take account of the censoring problem by eliminating all children younger than age 25.

Many of these children have parents who were very young when they were born. Because the parents' age at birth is likely related to observed and unobserved parental characteristics, censoring is no longer random. This means that Black *et al.* run the risk of introducing sample selection bias when they reduce their sample. With this form of non-random censoring, it is possible that their parental schooling effects estimated on the reduced sample in which younger parents are overrepresented are very different from the same parental schooling effects estimated on the full (but non-existing) sample without censored observations.

2.4. Intermediate School and Human Capital Outcomes

An alternative method to deal with censored observations is to work with multi-generational samples of parents with school-going children and consider intermediate outcomes of children that are both realized and available, such as birth weight (Currie and Moretti, 2003), test scores or grade repetition (Carneiro *et al.*, 2007; Oreopoulos *et al.*, 2006; Maurin and McNally, 2008). Without information on realized school outcomes of children, however, we do not really know how informative these intermediate outcomes are when it comes to assessing intergenerational schooling effects.

To formalize this argument, we let X_t be the intermediate outcome, which is generally realized during the children's compulsory schooling years. If we replace the children's censored and uncensored schooling observations with intermediate outcomes and run the following regression:

$$X_t = \alpha_0 + \alpha_1 S_{t-1} + v_t \quad (12)$$

we get an estimate of α_1 which measures the association between the schooling of the parent and the intermediate outcome of the child. The estimation of α_1 is interesting in its own right but not necessarily informative about β_1 . To let parameter α_1 be informative about β_1 , two conditions have to be met: (a) the children's intermediate school outcomes should be related to their realized years of schooling; and (b) the schooling of parents should not have an impact on children's schooling, conditional on the intermediate outcome.

Finding an intermediate outcome that fulfills these conditions is not easy. In fact, to the extent that similar performing children are treated differently in ways related to their parents' schooling, we may question whether most of the intermediate outcomes that are in use are informative about intergenerational schooling effects. If, for example, higher schooled parents find it easier to provide additional tutoring to young children who repeat a grade or receive low test scores than lower schooled parents, it is possible that the corresponding intermediate schooling estimates will not capture the parental treatment effects that children receive beyond their compulsory schooling years, and therefore miss the true impact of parental schooling on child schooling (Currie and Thomas, 2001; Erikson *et al.*, 2005).

While we think it is useful to actually test how informative intermediate schooling outcomes are, we must leave this for future research. The present dataset does not contain information on birth weight, test scores or grade repetition.

3. DATA

Our main analysis employs the WLS of 10,317 randomly sampled graduates from Wisconsin high schools in 1957. After the initial wave of data collection, primary respondents were re-interviewed in 1975, 1992 and 2004. Together with their parents' interview of 1964, these waves provide information on, among others, educational attainment of the original graduates, their parents and children. The original sample is broadly representative for white men and women, who have completed at least 12 years of schooling. For more detailed information on the WLS we refer to Wollmering (2007) and references therein.

In this paper we use all three waves and exploit those questions that are targeted at the educational attainment of the respondents' children. In 1975 children are still in school and parents are asked to express their expectations about their child's schooling. In particular, parents are asked how far they think their children will go in school.⁵ In 1992 most children left school, but some children were still in school. Information is collected on the highest grade of regular school ever attended; whether the highest grade is completed or not; and whether the highest grade is obtained during the survey year. In addition, respondents are asked whether their child completed the grade or year and whether their child attended a regular school (elementary, secondary, college and university) in the past 12 months. In 2004 all children finished their education, and respondents are asked to update their information regarding their children's completed schooling.

Our sample includes married respondents with children, who are observed in the three years 1975, 1992 and 2004. In 2004 information is gathered from 7265 of the 10,317 original respondents, of whom 5660 are married and have children in 1992. Of these 5660 respondents, 316 drop out because relevant schooling information for themselves or their spouses is missing. In 1975 respondents are asked to express their school expectations for only one of their children. This child is randomly selected by the interviewer. Of the randomly selected children, we select only those for whom information is collected in 1975, 1992 and 2004. This leaves us with a final sample consisting of 4097 parents and own birth children.⁶ Note that the selected children are on average 10 years old when the parent forms expectations regarding the child's schooling. Expectations are elicited from the originally sampled graduates, which include both men and women, so the expectations are in some cases formed by the father and in other cases by the mother. Summary statistics appear in Table I.

⁵ Parental expectations are expressed in levels. We convert levels into years in a similar way as Antonovics and Goldberger (2005, p. 1739) recode levels into years of schooling: less than high school . . . 10; high school graduate . . . 12; technical and vocational education . . . 13; some college . . . 14; college graduate . . . 16; MA or MS degree . . . 18; law degree, MD, DDS, DVM degree . . . 19; PhD. . . 20.

⁶ For some children who finished schooling in 1992, reported years of schooling in 2004 differ from years of schooling reported in 1992. For these observations we replace reported schooling in 1992 and 2004 by the maximum of the two. This is done for 208 children.

INTERGENERATIONAL MOBILITY AND CENSORING

Table I. Means and standard deviations of selected variables

	Mean	Std. Dev.
Completed years of schooling (2004)	14.50	2.30
Observed years of schooling (1992)	14.17	2.18
Expected years of schooling (1975)	14.78	1.93
Years of schooling mother	12.92	1.72
Years of schooling father	13.62	2.67
Observation censored in 1992	0.22	0.41
Gender (daughter)	0.49	0.50
Age (1992)	26.92	3.97
<i>N</i>	4097	

RESULTS

Table II presents estimates that come from our child–parent schooling regressions run on uncensored and censored samples of own birth children and their parents. All regressions include individual controls for the child’s age measured in years and gender. These parameters are not reported.

In the first panel the first three columns report estimates using the completed school measures as recorded in the 2004 sample. These estimates will serve as the baseline throughout the paper. In columns (1) and (2) the mother’s and father’s schooling measures are included as separate regressors. The coefficients on maternal and paternal years of schooling are equal to 0.45 and 0.35, respectively. These estimates are fully in line with those commonly found in the literature (Haveman and Wolfe, 1995): more schooled parents have more schooled children, and more schooled mothers matter more than more schooled fathers, at least for the school outcomes of children. Note that the parental associations as estimated not only include the influence from the given parent but also the influence from the spouse, which is due to assortative mating and the ensuing correlation between the parents’ schooling, or between everything else that correlates with schooling. In column (3) the mother’s and father’s schooling measures are therefore included simultaneously to control for assortative mating effects. We still find that more schooled parents get more schooled children, but that fathers and mothers now contribute equally to their offspring.

In the second three columns we estimate the same three equations using the observed school measures as recorded in the 1992 sample. With data that are partly censored we find, as expected, that all parental schooling estimates fall, but not by much. The last three columns, in which we express the difference between the child–parent estimates run on the censored and uncensored samples, indicate that the downward bias caused by the censoring is statistically significant and varies between 0.02 and 0.03.⁷

In the next three panels we report the estimates using alternative approaches to tackle the censoring problem: replacement of observed with expected years of schooling, maximum likelihood approach, and the elimination of all school-aged children. We find that the corrections do not affect our results qualitatively. In all three panels the estimates reported in columns (4), (5) and (6) show that more schooled parents get more schooled children and that mothers only matter more when

⁷ The previous schooling models are estimated combining both WLS samples where all coefficients vary by sample status. The interacted schooling estimate represents the absolute difference between mobility parameters.

Table II. Estimates of the effects of mother's and father's schooling on children's schooling

	Estimates without censoring			Estimates with censoring			Estimated differences ^a		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's schooling	0.45 <i>0.02***</i>		0.24 <i>0.02***</i>	0.42 <i>0.02***</i>		0.22 <i>0.02***</i>	-0.03 <i>0.01***</i>		-0.02 <i>0.01</i>
Father's schooling		0.35 <i>0.01***</i>	0.27 <i>0.01***</i>		0.32 <i>0.01***</i>	0.25 <i>0.01***</i>		-0.03 <i>0.01***</i>	-0.02 <i>0.01***</i>
<i>N</i>		4,097			4,097				
<i>N</i> ^c		0			874				
<i>Censored observations replaced with parental expectations^b</i>									
Mother's schooling				0.45 <i>0.02***</i>		0.23 <i>0.02***</i>	-0.00 <i>0.01</i>		-0.01 <i>0.01</i>
Father's schooling					0.35 <i>0.01***</i>	0.27 <i>0.01***</i>		0.00 <i>0.01</i>	0.00 <i>0.01</i>
<i>N</i>					4,097				
<i>N</i> ^c					874				
<i>Censored regression model</i>									
Mother's schooling				0.53 <i>0.02***</i>		0.30 <i>0.03***</i>	0.08 <i>0.01***</i>		0.05 <i>0.01***</i>
Father's schooling					0.39 <i>0.01***</i>	0.30 <i>0.02***</i>		0.04 <i>0.01***</i>	0.03 <i>0.01***</i>
<i>N</i>					4,097				
<i>N</i> ^c					874				
<i>Excluding all children younger than 25</i>									
Mother's schooling				0.50 <i>0.03***</i>		0.28 <i>0.03***</i>	0.05 <i>0.02***</i>		0.04 <i>0.01***</i>
Father's schooling					0.36 <i>0.02***</i>	0.28 <i>0.02***</i>		0.01 <i>0.01</i>	0.01 <i>0.01</i>
<i>N</i>					2,990				
<i>N</i> ^c					167				

All regressions include additional controls or the child's age and gender.

Standard errors are in italics; significant at: * 10% level; ** 5% level; *** 1% level.

N^c is the number of censored observations.

^a Estimates come from previous school models using censored and uncensored samples where all coefficients vary by sample status. The interacted schooling estimates represent differences between mobility parameters. Insignificance suggests the absence of structural differences.

^b Parental expectations are expressed in levels. We convert level into years as follows: less than high school ... 10; high school graduate ... 12; technical and vocational education ... 13; some college ... 14; college graduate ... 16; MA or MS degree ... 18; law degree, MD, DDS, DVM degree ... 19; PhD ... 20.

parental schooling estimates include assortative mating effects. But we do find that most of the corrections affect our results statistically. When compared to the uncorrected regression results using the censored sample, all three approaches remove the downward bias and give us—as they should—higher persistency estimates. When compared to those estimates obtained using the ultimate uncensored sample, the estimated differences in columns (7), (8) and (9) are relatively small in magnitude. Of the three correction methods, we find that the maximum likelihood and elimination approaches lead to estimates that are a bit too high, at least in most specifications. With these two corrections, the medicine appears to be no better than the malady. The approach to

treat parental expectations for young children as if they were realizations of completed schooling, however, works better. In fact, the replacement method fully eliminates the censoring bias, which we now precisely estimate at zero.

5. CAN WE TREAT EXPECTATIONS AS REALIZATIONS?

Our results in Table II suggest that parental expectations fix the censoring problem quite well. This is by no means a trivial result. After all, parents (in the WLS) form their expectations when their children are relatively young, about 10 to 15 years prior to school completion. In a recent paper Antonovics and Goldberger (2005, p. 1739) express their doubts regarding this particular correction method. We therefore perform additional robustness checks to see how sensitive the parental expectations solution is to a number of potential threats: the number of censored observations, prediction quality, and sample selection.

5.1. The Degree of Censoring

Our first concern is that the expectation method might work because the number of censored observations in our sample is relatively small. In Section 2 we showed that the bias introduced by replacing censored observations with parental expectations depends on the association between parental schooling and a combination of parental prediction error (η_t) and the degree of censoring (d_t); that is:

$$\text{plim } \tilde{\beta}_{1LS} - \beta_1 = \text{cov}(d_t \eta_t, S_{t-1}) / \text{var}(S_{t-1})$$

This is an expression we can actually test: least squares estimation of the regression of $d_t \eta_t$ on parental schooling. To see whether our results are sensitive to the number of censored observations, we estimate the bias of the replacement method on samples where we gradually increase the number of censored observations. We do this by calculating how many children would still be in school if we had observed them some years before 1992. For example, if a mother, who reports in 1992 that her child, born in 1967, completed 15 years of schooling, were interviewed in 1984 we recode the same child as being censored, assuming he/she left school in 1988 ($1967 + 6 + 15$). In 1984 the same mother would have reported that her child had 11 years of schooling, assuming that children start school at age 6 and have uninterrupted school careers.

The first panel of Table III contains the estimates of the bias when using the replacement method, for increasing numbers of censored observations, with additional controls for age and gender of the child. Up to censoring percentages of 60, we find that all the bias estimates are statistically insignificant and virtually zero, confirming our baseline result that the replacement method yields consistent persistency estimates. Up to censoring percentages of 90, the bias is negative but small, and often statistically insignificant. The procedure to replace the censored observations with expectations is statistically rejected, but only at the margin. Only when the percentage of censored observations becomes very large does the corresponding method to adjust for censoring fail. The slopes are negative and statistically significant. Would we fully rely on parental predictions, the implication is that the corresponding intergenerational schooling estimates are biased downwards. The negative bias further suggests that expectations regress to the mean faster than realizations do.

Table III. Estimating the bias for increasing censoring percentages

	Samples with increasing number of censored observations							
	20–30%	30–40%	40–50%	50–60%	60–70%	70–80%	80–90%	100%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Censored observations replaced with parental expectations</i>								
Mother's schooling	–0.01 <i>0.01</i>	–0.02 <i>0.02</i>	–0.02 <i>0.02</i>	–0.02 <i>0.02</i>	–0.02 <i>0.02</i>	–0.03 <i>0.02</i>	–0.04 <i>0.02</i>	–0.05 <i>0.02**</i>
Father's schooling	0.00 <i>0.01</i>	–0.01 <i>0.01</i>	–0.02 <i>0.01</i>	–0.02 <i>0.01</i>	–0.02 <i>0.01*</i>	–0.03 <i>0.01*</i>	–0.03 <i>0.01*</i>	–0.04 <i>0.02***</i>
<i>N</i>	4097	4097	4097	4097	4097	4097	4097	4097
<i>N</i> ^c	874	1285	1712	2280	2859	3157	3388	4097
<i>Censored regression model</i>								
Mother's schooling	0.05 <i>0.01***</i>	0.06 <i>0.02***</i>	0.06 <i>0.02***</i>	0.04 <i>0.03</i>	0.04 <i>0.03</i>	–0.01 <i>0.04</i>	–0.06 <i>0.04</i>	x
Father's schooling	0.03 <i>0.01***</i>	0.02 <i>0.01**</i>	0.01 <i>0.01</i>	0.00	–0.02 <i>0.02</i>	–0.06 <i>0.02***</i>	–0.09 <i>0.02***</i>	x
<i>N</i>	4097	4097	4097	4097	4097	4074	4006	x
<i>N</i> ^c	874	1285	1712	2280	2859	3134	3297	x
<i>Excluding all children younger than 25</i>								
Mother's schooling	0.04 <i>0.02**</i>	0.07 <i>0.04**</i>	0.03 <i>0.06</i>	0.02 <i>0.18</i>	x	x	x	x
Father's schooling	0.01 <i>0.01</i>	0.02 <i>0.02</i>	–0.01 <i>0.03</i>	–0.00 <i>0.06</i>	x	x	x	x
<i>N</i>	2990	1558	852	240	x	x	x	x

All regressions include additional controls or the child's age and gender. Standard errors are in italics; significant at: * 10% level; ** 5% level; *** 1% level. *N*^c is the number of censored observations.

In the second and third panels we also show results for the censored regression and elimination models. In the case of the censored regression approach, we find for small censoring percentages that the estimated bias is somewhat larger than the bias reported in the previous panel. When we increase censoring percentages, we find that the bias consistently falls. For censoring percentages around 50% the bias goes to zero and then becomes negative for samples where the majority of the children are still in school.⁸ When we exclude all children below 25 from our sample, we find for small percentages that the bias is positive, statistically significant and comparable to the bias of the censored regression model. Together with falling sample sizes, the bias declines for censoring percentages up to 60%. For samples where more than 60% of the observations are censored, all children are below the age of 25 and the elimination method no longer works. Overall, we believe that of the three different solutions to the censoring problem the replacement method appears to be the least sensitive to the number of censored observations.

⁸ This pattern is consistent with a bimodal schooling distribution. Arabmazar and Schmidt (1982) investigate the inconsistency of the related Tobit estimator as a consequence of different non-normal distributions. They find that the bias due to non-normality depends on the degree of censoring. They do not, however, investigate the consequences of a bimodal distribution. If we assume a bimodal distribution of years of schooling, our simulation results—not reported in the paper—bear out that the inconsistency of the maximum likelihood estimator is positive when about 25% of the observations are (right) censored, and negative when about 75% of the observations are censored.

5.2. Prediction Quality and Generalizability

Our second concern is that the replacement trick might work because of the non-representative nature of the WLS. The WLS only collects information on high school graduates and (because of that) systematically undersamples the lower educated individuals. If more schooled parents form more accurate expectations about their children's schooling, it is possible that our observation—the best approach is to replace censored observations with parental expectations—is driven by the sample design of the WLS, and does not hold in other datasets.⁹

To get an idea whether the mechanism of more schooled parents forming more accurate expectations is present among WLS parents, we first ask ourselves whether WLS parents can accurately predict their child's education. Figure 1 shows a histogram of the difference between parental expectations and realizations. Although for almost 40% of the children parental expectations coincide with realizations, there is quite some variation in how well parents can predict their child's schooling.

Figure 2 plots how the (absolute) difference between parental expectations and realizations varies with parental schooling, including the regression line that measures the relationship between (absolute) prediction error and parental schooling (controlling for age and gender of the child). In the top figures we see that the prediction error parents make is not fully independent of their amount of schooling. The slope coefficients (with standard error between brackets) which we estimate at -0.084 [0.021] for mother's schooling and -0.057 [0.013] for father's schooling indicate that higher schooled parents set their expectations a bit too low, while lower schooled parents set their expectations a bit too high. Since expectations that are too low can be as inaccurate as expectations that are too high, we also plot and estimate how prediction quality (measured as the absolute value of the prediction error) depends on parental schooling, age and gender of the child. The bottom figures show almost no visible relationship. The coefficients are 0.013 [0.017] for mother's schooling and -0.021 [0.011] for father's schooling, respectively. If we look at mothers, there appears to be no evidence that more schooled mothers make better predictions. If we look

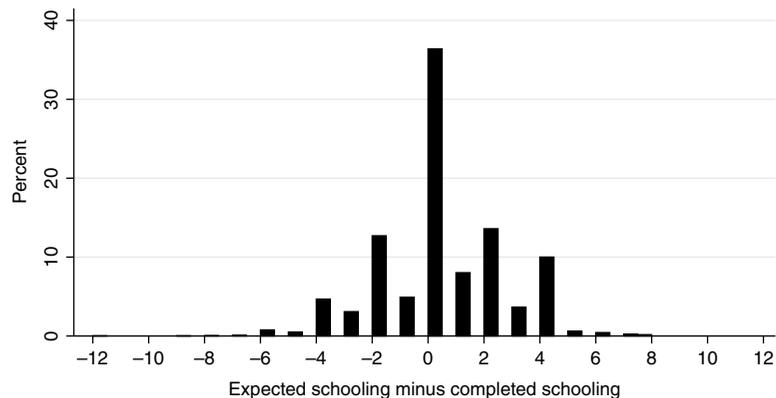


Figure 1. Difference between parental expectations and realized years of schooling

⁹ If more schooled respondents with more precise expectations about their children's schooling are more likely to answer the 1975, 1992 and 2004 questionnaires, sample attrition will tend to improve the performance of the replacement method relative to that of the other two competing methods.

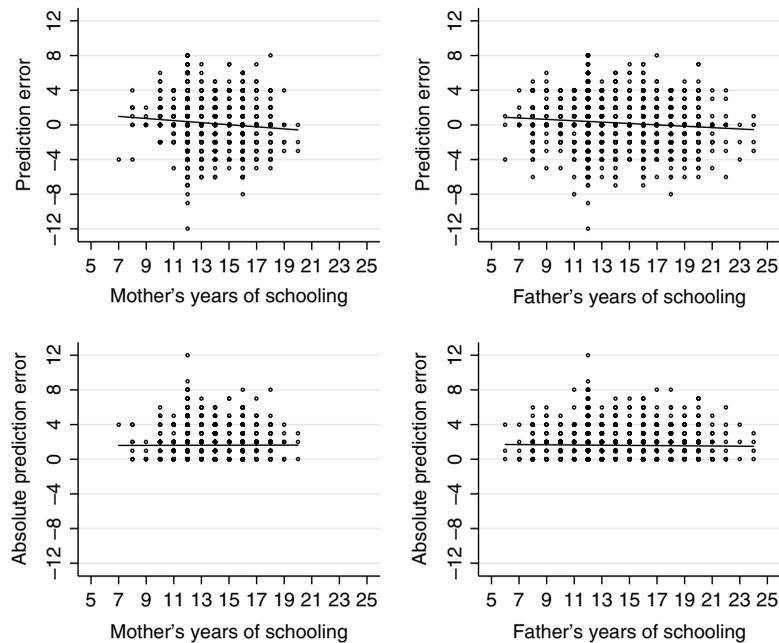


Figure 2. (Absolute) prediction errors and parental schooling

at fathers, we do find that prediction quality improves with years of schooling, but it is only at the margin. These weak correlations, we think, raise the generalizability of our findings.

To get an idea whether WLS parents form expectations about their children's schooling in similar ways as other parents do in other datasets, we compare our WLS results with results that come from two more familiar multigenerational US samples drawn from the Michigan Panel Study of Income Dynamics (PSID)¹⁰ and the National Longitudinal Survey of Youth (NLSY).¹¹ In all three samples we focus on mothers who expressed their expectations. The WLS and PSID samples are comparable with respect to the survey year expectations are elicited: all mothers were asked to form expectations in either 1975 or 1976. The WLS and NLSY samples are comparable with respect to the mother's age when expectations are elicited: all mothers are in their mid thirties when they report their expectations.

In the first panel of Table IV we report sample means and standard deviations for some of the variables we will study below. The summary statistics clearly illustrate that the mothers in our WLS sample are very different from the mothers in the PSID and NLSY samples. Compared

¹⁰ The PSID is a longitudinal survey that began to collect information on about 5000 families in 1968. We rely on the unbalanced version, in which low-income families are overrepresented. We use PSID data in 1976, when mothers with children younger than 13 (including children they were still expecting to have) were asked how much schooling they expected their children to have. We focus on two-parent families. Our sample consists of 1936 mothers for whom we know age, schooling, schooling of her spouse, and expectations regarding her children's schooling.

¹¹ The NLSY study is a representative longitudinal panel. The initial sample started in 1979 and included young men and women aged between 14 and 22. When these women became mothers, the NLSY began to collect information on children as well, including information on expected schooling for all children between 5 and 15 years old. We use the 1996 survey as the comparison year when mothers are between 31 and 39 years old. We limit our sample to children who were living with both parents. Our sample consists of 1587 mothers with 2635 children.

INTERGENERATIONAL MOBILITY AND CENSORING

to the PSID mothers, we observe that WLS mothers are much better educated. This result is perhaps not surprising given that we use the unbalanced version of the PSID sample in which low income families are oversampled. Compared to a representative sample of NLSY mothers, we find that WLS mothers have similar amounts of schooling. Despite these similarities, this result indicates that more schooled mothers are clearly oversampled in our WLS sample given that average schooling has increased over time, and that WLS mothers are almost 20 years older than NLSY mothers.

In the second panel of Table IV we compare estimates that come from regressions of parental expectations on years of schooling of the mother, years of schooling of the father, and the mother's age (when information on expectations is collected) using the three different samples. In the first three columns we show results using our WLS sample. We find that more schooled mothers with more schooled spouses expect their children to do better in school. In the next six columns of Table IV we report the parental schooling estimates from the two other samples. Compared to the parental expectation results using WLS mothers, we find that the estimates obtained with the PSID and NLSY samples are quite similar. These findings imply that mothers in the WLS form expectations about their children's completed schooling in similar ways as mothers do in the PSID or the NLSY. But more importantly, these findings also imply that the replacement method will probably also work well in other more (or less) representative datasets.

Table IV. Regressing parental expectations on parents' schooling using samples of the WLS, PSID and NLSY

Sample statistics	WLS		PSID		NLSY	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Expected schooling	14.66 ^a	<i>1.93</i>	14.40 ^b	<i>1.88</i>	15.42 ^c	<i>2.13</i>
Mother's schooling	12.93	<i>1.65</i>	12.00	<i>2.28</i>	12.83	<i>2.40</i>
Father's schooling	13.38	<i>2.82</i>	12.06	<i>2.91</i>	12.97	<i>2.59</i>
Age mother	36.08	<i>0.44</i>	29.47	<i>8.29</i>	35.00	<i>2.24</i>
Age child	10.68	<i>3.87</i>			9.65	<i>2.70</i>
<i>N</i>	2225		1936		2635	
<i>Regressing expected schooling on mother's and father's schooling</i>						
Mother's schooling	0.39	<i>0.19</i>	0.37	<i>0.23</i>	0.31	<i>0.19</i>
	<i>0.02***</i>	<i>0.03***</i>	<i>0.02***</i>	<i>0.02***</i>	<i>0.02***</i>	<i>0.03***</i>
Father's schooling		0.29	0.23	0.30	0.19	0.30
		<i>0.01***</i>	<i>0.02***</i>	<i>0.01***</i>	<i>0.02***</i>	<i>0.02***</i>

^a Parents are asked: 'How far do you think your child probably will go in school?' Expectations reported in 1975 by mother in levels. We convert levels into years as follows: less than high school ... 10; high school graduate ... 12; technical and vocational education ... 13; some college ... 14; college graduate ... 16; MA or MS degree ... 18; law degree, MD, DDS, DVM degree ... 19; PhD ... 20.

^b Mothers are asked: 'About how much education do you think your children will have when they stop going to school?' Expectations reported in 1976 for children below 13 and/or future children in six levels: some high school ... 10; high school ... 12; high school plus non-academic training ... 13; college but no degree ... 14; college BA and no advanced degree ... 16; college and advanced or professional degree ... 19.

^c Mothers are asked: 'Looking ahead, how far do you think your child will go in school?' Expectations reported in 1996 for each child between 5 and 15 years old in five levels: leave high school before graduation ... 10; graduate from high school ... 12; get some college or other training ... 14; graduate from college ... 16; take further training after college ... 19.

In the second panel all regressions include additional controls for age of the mother at date of the interview. In the second panel standard errors are in italics; significant at: * 10% level; ** 5% level; *** 1% level.

5.3. Patience versus Impatience

Since the replacement method is not as disconcerting as Antonovics and Goldberger say it is, it is interesting to see what happens if a researcher is very impatient and wants to estimate the degree of intergenerational mobility when none of the children has finished their schooling. In Table IV we report those estimates researchers would get had they relied exclusively on the 1975 sample and replaced all observations by parental expectations.¹² In the WLS specification, where we control for the schooling of both parents, we find child–parent associations of 0.19 and 0.23 for mothers and fathers, respectively. Compared to the associations of 0.24 and 0.27 obtained with the uncensored WLS sample, we see that the WLS estimates in Table IV are statistically but not substantially different, which is quite remarkable given that schooling expectations were measured when almost all children were still in primary school.

5.4. Maximum Likelihood and the Elimination Approach

So far, the sensitivity analysis has concentrated on mechanisms that could possibly invalidate the method to replace censored school observations with parental expectations. But perhaps it is also informative to understand why the maximum likelihood and elimination approach produce biased mobility estimates, even though the bias as reported in Table II is not substantial.

We begin with the maximum likelihood approach. One likely candidate to explain the upward bias of the maximum likelihood approach would be a normality violation. It is unlikely that schooling is normally distributed—the more appropriate distribution of the child’s completed education is bimodal with peaks around 12 and 16 years (see also footnote 8). Another possibility is that heteroskedasticity is causing the inconsistent estimates. Using the uncensored 2004 sample we can test whether the normality and or homoskedasticity assumptions are violated. Our results show that the null hypotheses of normality and homoskedasticity are both rejected.¹³

A candidate to explain the inconsistencies caused by the elimination procedure would relate to the fact that by eliminating all children below 25 we are left with a sample consisting of parents who chose to have children at a relatively early age. These parents are likely different in both observed and unobserved characteristics which can cause the upward bias we observe in Table II. Parents who choose to have children at an early age, for example, are mostly lower educated. If mobility is lower at the lower end of the distribution (Oreopoulos *et al.*, 2006) the elimination of mostly children from higher educated parents would lead to an estimate of the mobility parameter that is too high.

6. CONCLUDING REMARKS

Recent studies on intergenerational schooling mobility often rely on samples in which information on the child’s completed schooling is not always available. Unfortunately, solutions offered to handle censored samples do not always work, and should be further scrutinized.

¹² Note that the other two methods do not work with samples where all the observations are censored.

¹³ The tests for normality and homoskedasticity are performed on the specification including both mother’s and father’s schooling as regressors. The *p*-value of the Breusch–Pagan/Cook–Weisberg test is equal to 0.007, and the null hypothesis of homoskedasticity is rejected. The *p*-value of the skewness/kurtosis tests for normality is equal to 0.000, and the null hypothesis of normality is therefore also rejected.

INTERGENERATIONAL MOBILITY AND CENSORING

This is what we do in this paper. We first estimate the impact of mother's and father's schooling on child's schooling using censored and uncensored samples, and investigate the consequences of three different methods that deal with censored observations: replacement of observed with expected years of schooling, maximum likelihood approach, and elimination of all school-aged children.

Our basic result is that parental schooling effects fall, but not by much, when intergenerational mobility models are estimated on censored samples and rise, again not by much, when censored observations are tackled by any of the three correction methods. Of the three methods, the one that treats parental expectations as if they were realizations performs best.

This replacement result depends, however, on the degree of censoring. For samples that are largely incomplete the method does give a small (negative) bias. Nonetheless, it does not matter (much) whether researchers are patient or impatient: whether we fully rely on parental expectations, or whether we use realizations measured 30 years later, the mobility estimates are not substantially different.

An important question is whether our replacement result in this paper provides some guidance as to whether it might work in other datasets as well. We are quite confident it does, for two reasons. First, it is possible to apply the replacement method to other datasets. There are other, more familiar, intergenerational data sources available that collect information on expected schooling when children are still in school. Second, we find that parents in the WLS form expectations about their children's completed schooling in similar ways as many other parents do. It seems therefore reasonable to speculate that our replacement results will persist in other intergenerational data sources as well. We think, however, that further insights on these remarkable findings can be gained by more research investigating how parents actually form their expectations.

ACKNOWLEDGEMENTS

The authors thank Reyn van Ewijk, Hessel Oosterbeek and participants in seminars/workshops at the Tinbergen Institute (Amsterdam, 2006), University of Oxford (Oxford, 2006), University of Essex (Essex, 2007), CEMFI (Madrid, 2007), IZA (2007) and LEaF conference (London, 2007), as well as three anonymous referees for their helpful comments. They further thank the Dutch National Science Foundation for its support under VIDI Grant 452.03.309. Support for collection and dissemination of data from the Wisconsin Longitudinal Study has been provided by the National Institute on Aging (AG-9775), the National Science Foundation (SBR-9320660), the Spencer Foundation, and the Center for Demography and Ecology and the Vilas Estate Trust at the University of Wisconsin–Madison. Data and documentation from the Wisconsin Longitudinal Study are available at <http://www.ssc.wisc.edu/wlsresearch/documentation>.

REFERENCES

- Antonovics KL, Goldberger AS. 2005. Does increasing women's schooling raise the schooling of the next generation? Comment. *American Economic Review* **95**: 1738–1744.
- Arabmazar A, Schmidt P. 1982. An investigation of the robustness of the tobit estimator to non-normality. *Econometrica* **50**: 1055–1064. DOI:10.2307/1912776.
- Behrman JR, Rosenzweig MR. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review* **92**: 323–334.

- Behrman JR, Rosenzweig MR. 2005. Does increasing women's schooling raise the schooling of the next generation? Reply. *American Economic Review* **95**: 1745–1751.
- Black SE, Devereux PJ, Salvanes KG. 2005. Why the apple doesn't fall far: understanding intergenerational transmission of human capital. *American Economic Review* **95**: 437–449.
- Böhlmark A, Lindquist MJ. 2006. Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. *Journal of Labor Economics* **24**: 879–896.
- Carneiro P, Meghir C, Patey M. 2007. *Maternal education, home environments and the development of children and adolescents*. IZA Working Paper No. 3072.
- Chevalier A. 2004. *Parental education and child's education: a natural experiment*. IZA Working Paper No. 1153.
- Currie J, Moretti E. 2003. Mother's education and the intergenerational transmission of human capital: evidence from college openings. *Quarterly Journal of Economics* **118**: 1495–1532.
- Currie J, Thomas D. 2001. Early test scores, school quality and SES: longrun effects on wage and employment outcomes. *Research in Labor Economics* **20**: 103–132.
- Erikson R, Goldthorpe JH, Jackson M, Yaish M, Cox DR. 2005. On class differentials in educational attainment. *Proceedings of the National Academies of Sciences* **102**: 9730–9733.
- Haider SJ, Solon G. 2006. Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* **96**: 1308–1320.
- Haveman R, Wolfe B. 1995. The determinants of children attainments: a review of methods and findings. *Journal of Economic Literature* **33**: 1829–1878.
- Maurin E, McNally S. 2008. Vive la révolution! Long term returns of 1968 to the angry students. *Journal of Labor Economics* **26**: 1–33.
- Oreopoulos P, Page MEJ, Stevens AH. 2006. The intergenerational effects of compulsory schooling. *Journal of Labor Economics* **24**: 729–760.
- Plug E. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *American Economic Review* **94**: 358–368.
- Wisconsin Longitudinal Study (WLS). 2004. Version 11.0 Preliminary release. University of Wisconsin–Madison, Madison, WI. <http://www.ssc.wisc.edu/wls/documentation/> [3 August 2009].
- Wollmering E (eds). 2007. *Wisconsin Longitudinal Study Handbook* (12.10.07) with contributions by Wisconsin Longitudinal Study Staff.