Family background and children's schooling outcomes

de Haan, M.

Chapter 2

Estimating intergenerational schooling mobility on censored samples: Consequences and remedies[1]

2.1 Introduction

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

$$Y_t = \alpha + \beta Y_{t-1} + u_t \tag{2.1}$$

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 which measures the outcome association between parent and child. Estimating β, however, puts strong requirements on data. In household surveys, in particular, the collection of information on realized outcomes of children is often problematic. If Y represents (permanent) income, for example, income information of children is rarely available. Most children who still live with their parents do not work. And even if information is available, intergenerational mobility estimates will be biased downwards when the children's income is measured too early in life (Haider and Solon (2006); Bölmark and Lindquist (2006)).

[1] Joint work with Erik Plug

In this chapter 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 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 mobility estimates that are too low. Second, we observe that censoring is a widely spread phenomenon. Of the recent studies that (aim to) make a distinction between causation and selection, almost all rely on samples with incomplete information on adult children. Among these studies are Behrman and Rosenzweig (2002); Chevalier (2004); Currie and Moretti (2003); Plug (2004); Black *et al*. (2005A); Carneiro *et al*. (2007); Oreopoulos *et al*. (2006); Maurin and McNally (2008). And third, the solutions offered to handle censored samples rely on assumptions that may not hold in practice, resulting in biased mobility estimates.

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 wouldn't 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. Since the latter approach certainly merits serious consideration, it is important to know (more) about how the available correction methods deal with censored observations.

Three correction methods are currently in use: maximum likelihood approach, replacement of observed with expected years of schooling, and elimination of all school-aged children.[2] In this chapter we apply these three different methods to correct for the censor-

[2] An alternative method to deal with censored observations is to look at intermediate outcomes that are realized and available, such as birth weight (Currie and Moretti 2004), grade repetition (Carneiro *et al*. 2007; Oreopoulos *et al*. 2006; Maurin and McNally 2008) or post-compulsory schooling attendance (Chevalier 2004). Without information on realized school outcomes of children, however, we do not know how informative these intermediate outcomes are when it comes to assessing intergenerational schooling effects.

ing problem to one particular data set: 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 still in school: the sample includes information on expected schooling. In 1992 the same children are about to complete school: the sample is a censored sample. In 2004 all children have passed their school-going age: the sample contains information on completed schooling.

Our contributions are twofold. First, we present new estimates of the intergenerational mobility of schooling. In addition, we make a distinction between own birth children and adoptees to investigate how much inherited abilities contribute to the impact of parental schooling. With updated 2004 samples, we estimate the ultimate mobility models in which censored observations are absent. And 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 chapter continues as follows. Section 2.2 models the intergenerational mobility of schooling, focuses on the problem that children who are still in school generate censored observations, and provides some intuition of the various solutions to it. Section 2.3 provides a brief description of the data. Section 2.4 presents and compares the parameter estimates. Section 2.5 evaluates the correction methods and presents a number of robustness tests and Section 2.6 concludes.

2.2 Mobility models using censored data

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

$$S_t = \alpha + \beta S_{t-1} + \varepsilon_t, \tag{2.2}$$

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 mobility parameter β measures the association between the schooling of parent and child. With information on S_t and S_{t-1}, the least-squares estimator is defined as

$$\text{plim } \beta_{LS} = \text{cov}(S_t, S_{t-1})/\text{var}(S_{t-1}) = \beta. \tag{2.3}$$

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, the intergenerational schooling model needs to be rewritten as

$$S_t^c = \begin{cases} S_t^c = S_t \text{ if } d_t = 0, \\ S_t^c < S_t \text{ if } d_t = 1, \end{cases} \tag{2.4}$$

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 if it were their completed years, the estimation of S_t^c on S_{t-1} using ordinary least squares gives us a mobility parameter that is too low. The intuition is as follows. We know that (*a*) more schooled children (with more schooled parents) are more likely to be censored; and (*b*) observed years are smaller than or equal to the completed 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 now apply least squares to estimate the model

$$S_t^c = \alpha^c + \beta^c S_{t-1} + \varepsilon_t^c, \tag{2.5}$$

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

$$\text{plim}\, \beta_{LS}^c = \text{cov}(S_t^c, S_{t-1})/\text{var}(S_{t-1}) \leq \text{cov}(S_t, S_{t-1})/\text{var}(S_{t-1}) = \beta. \tag{2.6}$$

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

A censored regression model

Plug (2004) exploits the 1992 wave of the WLS to estimate the effect of fathers and mothers schooling on child's schooling using samples of biological and adopted children. In 1992, however, many 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 homoskedastic errors the likelihood function is

$$L(\theta) = \prod_{i=1}^{N} [\phi(S_t \mid S_{t-1}, \theta)]^{1-d_t} [1 - \Phi(S_t^c \mid S_{t-1}, \theta)]^{d_t}, \tag{2.7}$$

where ϕ and Φ represent normal density and distribution functions, θ are the distribution parameters that include β, and i indexes the family in which the child is born and raised. Maximization of (2.7) yields a consistent estimator of β, unless the error distribution is incorrectly specified, being non-normally distributed or having heteroskedastic errors of unknown form.

Eliminating all school-going aged children

Black *et al.* (2005A) estimate the effect of parental schooling on child schooling using a reform in compulsory schooling in Norway during the sixties and early seventies 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 finished their schooling yet by the time they appear in their sample. They take account of the censoring problem by eliminating all children younger than age 25.

Many of these children are likely to have parents who were very young when they were born. Black *et al.* therefore run the risk of introducing sample selection bias when they reduce their sample. The argument is that censoring is not random but related to observed and unobserved parental characteristics, and that the corresponding estimate of the effect of parental schooling using the reduced sample can be biased.

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 completed their schooling yet; this is the case for more than 50% of their sample.[3]

Behrman and Rosenzweig propose to replace their censored observations with parental expectations and treat these expectations as if they were school realizations for children

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

with unfinished schooling. This gives the following school variable for the child

$$\widetilde{S}_t = \begin{cases} S_t \text{ if } d_t = 0, \\ S_t^e \text{ if } d_t = 1, \end{cases} \tag{2.8}$$

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

$$S_t^e = S_t + \eta_t, \tag{2.9}$$

where η_t is the error parents make in predicting their child's completed schooling.[4] Combining (2.2), (2.8) and (2.9) leads to

$$\widetilde{S}_t = \alpha + \beta S_{t-1} + d_t \eta_t + \varepsilon_t. \tag{2.10}$$

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

$$\text{plim } \widetilde{\beta}_{LS} = \text{cov}(\widetilde{S}_t, S_{t-1})/\text{var}(S_{t-1}) = \beta + \text{cov}(d_t \eta_t, S_{t-1})/\text{var}(S_{t-1}). \tag{2.11}$$

Only if $\text{cov}(d_t \eta_t, S_{t-1})$ equals 0, Behrman and Rosenzweig's original solution produces an unbiased estimate of β. 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 on in this chapter.

[4] We 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 Data

Our analysis employs the Wisconsin Longitudinal Study (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 twelve years of schooling. For more detailed information on the WLS we refer to Sewell *et al.* (2004) and WLS (2006) and the references therein.

In this chapter we use all three waves and exploit those questions that are targeted at the educational attainment of the respondent's children. In 1975 children are still in school and questions are asked to elicit parental expectations.[5] In 1992 children are about to complete or just completed their schooling and 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, colleges, and universities) in the past 12 months. In 2004 these children all 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 years 1975, 1992 and 2004. In 2004 information is gathered from 7,265 of the 10,317 original respondents, of whom 5,630 are married and have children older than 12 in 1992. Of these 5,630 respondents 442 drop out because relevant schooling information of themselves, their spouses and children is missing.[6] This leaves us with a sample of 5,188 re-

[5]Parental expectations are expressed in levels. We convert levels into years in a similar way as Antonovics and Goldberger (2005, pp.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; M.A. or M.S. degree…18; Law degree, M.D., D.D.S., D.V.M. degree…19; Ph.D.…20.

[6]For some children who finished schooling in 1992, reported years of schooling in 2004 differs from years of schooling

spondents having 14,524 own birth children and 520 adopted children. Note that in 1975 respondents are asked to express their school expectations for only one of their children. This child is randomly selected by the interviewer. The censoring analysis in which we replace censored with expected school measures relies therefore on a much smaller sample, consisting of 4,097 own birth and 52 adopted children. 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 2.1.

Table 2.1: Means and standard deviations of selected variables in WLS samples

	Own birth children		Adoptees	
	Mean	Std. dev.	Mean	Std. dev.
Completed years of schooling (2004)	14.37	2.28	14.03	2.09
Observed years of schooling (1992)	13.84	2.34	13.25	2.12
Expected years of schooling[a] (1975)	14.78	1.93	15.08	1.97
Years of schooling mother	12.83	1.65	13.27	1.92
Years of schooling father	13.50	2.66	14.49	2.97
Observation censored in 1992	0.23	0.42	0.39	0.49
Gender (daughter)	0.50	0.50	0.49	0.50
Age (1992)	26.52	4.51	23.97	4.62
N	14,524		520	

[a]Parental expectations are asked for only one of the respondent's children. Means and standard deviations are therefore calculated on smaller samples of respectively 4,097 and 52 observations.

reported in 1992. For these observations we replace reported schooling in 1992 and 2004 by the maximum of the two. This is done for 487 own birth children and 39 adoptees.

2.4 Results

Table 2.2 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 and gender. These parameters are not reported.[7]

In the first three columns we report estimates using the completed school measures as recorded in the 2004 sample. In columns (1) and (2) the mother's and father's schooling measures are included as separate regressors. We find that more schooled parents have more schooled children, and that more schooled mothers matter more than more schooled fathers. In column (3) the mother's and father's schooling measures are 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. It is clear that these estimates are biased. The last three columns, in which we express the difference between mobility estimates run on the censored and uncensored samples, indicate that the downward bias caused by the censoring is statistically significant and varies between the 6 and 16 percent.[8]

In the next three panels we report the estimates using alternative approaches to tackle the censoring problem: maximum likelihood approach, elimination of all school-aged children, and replacement of observed with expected years of schooling.

[7]The estimations use all children, including all children raised in one family. With multiple family observations, standard errors are not independent within families and are biased downwards. We therefore estimate the model with clustered error terms to control for correlation within families.

[8]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 2.2: Estimates of the effects of mother's and father's schooling on own birth children's schooling

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's	0.46***		0.24***	0.41***		0.20***	-0.05***		-0.04***
schooling	(0.01)		(0.02)	(0.01)		(0.02)	(0.01)		(0.01)
Father's		0.34***	0.26***		0.31***	0.24***		-0.03***	-0.02***
schooling		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.01)
N		14,524			14,524				
N^c		0			3,278				
CENSORED REGRESSION MODEL									
Mother's				0.54***		0.29***	0.08***		0.05***
schooling				(0.02)		(0.02)	(0.01)		(0.01)
Father's					0.38***	0.29***		0.04***	0.03***
schooling					(0.01)	(0.01)		(0.00)	(0.00)
N					14,524				
N^c					3,278				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's				0.50***		0.26***	0.04***		0.03***
schooling				(0.02)		(0.02)	(0.01)		(0.01)
Father's					0.36***	0.28***		0.02***	0.02***
schooling					(0.01)	(0.01)		(0.01)	(0.01)
N					10,143				
N^c					508				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS[a]									
Mother's	0.45***		0.24***	0.45***		0.23***	-0.00		-0.01
schooling	(0.02)		(0.02)	(0.02)		(0.02)	(0.01)		(0.01)
Father's		0.35***	0.27***		0.35***	0.27***		0.00	0.00
schooling		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
N		4,097			4,097				
N^c		0			874				

All regressions include additional controls or the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.
[a]Samples are smaller because expectations are elicited for only one of the respondent's 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 parental schooling estimates include assortative mating effects. But we do find that the corrections affect our results quantitatively. 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 mobility estimates. When compared to those estimates obtained using the ultimate uncensored sample, the estimated differences in columns (7), (8) and (9) indicate that especially maximum likelihood and elimination approaches lead to mobility estimates that are too high. Instead of providing consistent estimates, these two censoring corrections cause an upward bias that is statistically significant and varies between the 6 and 21 percent. 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, does better. The bias is at most 4 percent and never statistically significant.

Adoption results

Recall that all the positive mobility estimates reported in Table 2.2 include the contribution of inherited abilities to intergenerational schooling transfers. To get rid of the effects caused by the parents' genes, we run our child-parent schooling regressions on samples of adoptees and their adoptive parents. This is done in Table 2.3 which has the same format as Table 2.2. We do this for two reasons. The first reason is that recent intergenerational mobility studies focus their attention on the parameter that measures the causal impact of the parent's schooling on that of the child. The second reason is that it is not obvious how censoring affects the schooling association between parent and child, net of the inherited endowments of parents.

Table 2.3: Estimates of the effects of mother's and father's schooling on adopted children's schooling

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's	0.22***		0.08	0.19***		0.06	-0.03		-0.02
schooling	(0.05)		(0.06)	(0.05)		(0.06)	(0.05)		(0.05)
Father's		0.22***	0.19***		0.19***	0.17***		-0.03	-0.02
schooling		(0.03)	(0.04)		(0.03)	(0.03)		(0.03)	(0.03)
N		520			520				
N^c		0			198				
CENSORED REGRESSION MODEL									
Mother's				0.21***		0.06	-0.01		-0.02
schooling				(0.07)		(0.07)	(0.03)		(0.04)
Father's					0.23***	0.21***		0.01	0.02
schooling					(0.04)	(0.04)		(0.02)	(0.02)
N					520				
N^c					198				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's				0.32***		0.19**	0.10		0.11
schooling				(0.08)		(0.09)	(0.07)		(0.08)
Father's					0.26***	0.20***		0.04	0.01
schooling					(0.05)	(0.06)		(0.05)	(0.05)
N					216				
N^c					19				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS[a]									
Mother's	0.13		–0.13	0.25		–0.02	0.12		0.11
schooling	*0.15*		*0.18*	*0.16*		*0.19*	*0.08*		*0.10*
Father's		0.23	0.28		0.28	0.28		0.05	0.00
schooling		*0.09***	*0.11***		*0.09****	*0.12***		*0.07*	*0.09*
N		52			52				
N^c					21				

All regressions include additional controls or the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.
[a]Samples are smaller because expectations are elicited for only one of the respondent's children.

In columns (1), (2) and (3) the results using the uncensored sample of adoptees are shown. We find that all the estimated effects of parental schooling drop when we move from own birth children to adoptees. This is consistent with the idea that part of the child's schooling is inherited. In column (3) we take the impact of the marriage partner into account, and find that the estimates fall only little for fathers, but much more for mothers. The maternal schooling effect reduces to 0.08 and lacks statistical significance, while the paternal schooling effect remains much larger in magnitude: 0.22 and 0.19 with or without taking into account the effect of his marriage partner. These findings are, as such, fully in line with those reported in Plug (2004) but also in Behrman and Rosenzweig (2002, 2005) and Björklund *et al.* (2006).

In columns (4), (5) and (6) of the first panel we see that the estimated effects remain qualitatively very similar, except that they are all smaller than the corresponding point estimates in the first three columns. This is not unexpected when we switch from the uncensored to the censored adoption sample. The bias is bigger than in our previous samples and varies between 11 and 21 percent. Probably because of the smaller samples, the censoring bias is rather imprecisely estimated, and never statistically significant (see columns (7), (8) and (9)).

In the remaining panels we evaluate the various solutions to the censoring bias. Compared to the uncorrected regression results using the censored adoption sample the three approaches produce (almost always) higher mobility estimates and thus appear to remove the downward bias. Compared to the results in Table 2.2 the estimated impacts of parental schooling drop and inherited abilities and assortative mating seem to play a more important role for mothers than for fathers. There is one exception. In the last panel where we eliminate all school-going aged children, we obtain a coefficient on mother's schooling that is statistically significant and almost of the same size as the coefficient on father's schooling. In columns (7), (8) and (9) we report on the differences between the estimates using the three approaches and the estimates without censoring. These differences are larger than those differences reported in Table 2.2 using the sample of own birth children. But since resulting differences are all statistically insignificant, it is difficult to draw firm conclusions about the validity of each solution.

2.5 Can we treat expectations as realizations?

Our results in Table 2.2 suggest that parental expectations fix the censoring problem quite well.[9] This is by no means a trivial result. In a recent paper Antonovics and Goldberger (2005 p.1739) express their doubt 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, sample selection and prediction quality.

The degree of censoring

Our first concern is that the expectation method might work well because the number of censored observations in our sample is relatively small. In Section 2.2 we showed that the bias introduced by replacing censored observations with parental expectations depends on the correlation of parents' schooling with parental prediction errors (η_t) and the degree of censoring (d_t); that is,

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

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

[9]In this Section we restrict our attention to the own birth results. Since the bias estimates in our adoption analysis are too imprecise and not informative about the preferred correction approach, we have decided to ignore the adoption results, at least, when we evaluate the three correction methods.

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 2.4 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 mobility 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, the corresponding method to adjust for censoring fails. The slopes are negative and statistically significant. Would we fully rely on parental predictions, the implication is that the corresponding intergenerational mobility estimates are biased downwards. The negative bias further suggests that expectations regress to the mean faster than realizations do.

In the second and third panel we also show results for the censored regression and elimination models. In 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 percent the bias goes to zero and then becomes negative for samples where the majority of the children is still in school.[10]

[10]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, however, not investigate the consequences of a bimodal distribution. If we assume a bimodal distribution of years of schooling, our simulation results –not reported in the chapter– bear out

Table 2.4: Estimating the bias for increasing censoring percentages

	Samples with increasing numbers 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)
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS								
Mother's schooling	-0.01	-0.02	-0.02	-0.02	-0.02	-0.03	-0.04	-0.05**
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Father's schooling	0.00	-0.01	-0.02	-0.02	-0.02*	-0.03*	-0.03*	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
N	4097	4097	4097	4097	4097	4097	4097	4097
N^c	874[a]	1285	1712	2280	2859	3157	3388	4097
CENSORED REGRESSION MODEL								
Mother's schooling	0.05***	0.06***	0.06***	0.04	0.04	-0.01	-0.06	x
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	
Father's schooling	0.03***	0.02**	0.01	0.00	-0.02	-0.06***	-0.09***	x
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	
N	4097	4097	4097	4097	4097	4074	4006	x
N^c	874	1285	1712	2280	2859	3134	3297	x
EXCLUDING ALL CHILDREN YOUNGER THAN 25								
Mother's schooling	0.04**	0.07**	0.03	0.02	x	x	x	x
	(0.02)	(0.04)	(0.06)	(0.18)				
Father's schooling	0.01	0.02	-0.01	-0.00	x	x	x	x
	(0.01)	(0.02)	(0.03)	(0.06)				
N	2990	1558	852	240	x	x	x	x

All regressions include controls for the child's age and gender. Standard errors are in parentheses.* significant at 10% level, ** significant at 5% level, *** significant at 1% level.[a]Number corresponds to true number of censored observations in 1992.

that the inconsistency of the maximum likelihood estimator is positive when about 25 percent of the observations is (right) censored, and negative when about 75 percent of the observations is censored.

In case 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 percent. For samples where more than 60 percent of the observations is 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.

Prediction quality

Our second concern is that the replacement trick might work because of the nonrepresentative nature of the WLS. The WLS only collects information on high school graduates and (because of that) systematically under-samples 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 data sets.

Other data sources with panel information on completed schooling and expected schooling (when the same children were still in school) would resolve part of this external validity discussion. These two measures, however, are rarely collected in large systematic data sets. It is still possible 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 perfectly predict their child's education. Figure 2.1 shows a histogram of the difference between parental expectations and realizations.

Although for more than 35 percent of the children parental expectations coincide with realizations, there is quite some variation in how well parents can predict their child's schooling. To explore this variation further, Figure 2.2 shows scatter diagrams and lowess estimation of the relation between the prediction error and mother's and father's schooling. This figure indicates that there is no strong relation between parental schooling and the difference between expected and completed schooling of the child. We also estimate how the absolute prediction error depends on parental schooling, age and gender of the child. We observe that the coefficient (with corresponding absolute *t* ratio between brackets) is 0.013 [0.76] for mother's schooling and -0.021 [1.88] for father's schooling. If we look at mothers, there appears to be no evidence that more schooled mothers make better predictions. If we look 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.

Figure 2.1: Difference between parental expectations and completed years of schooling

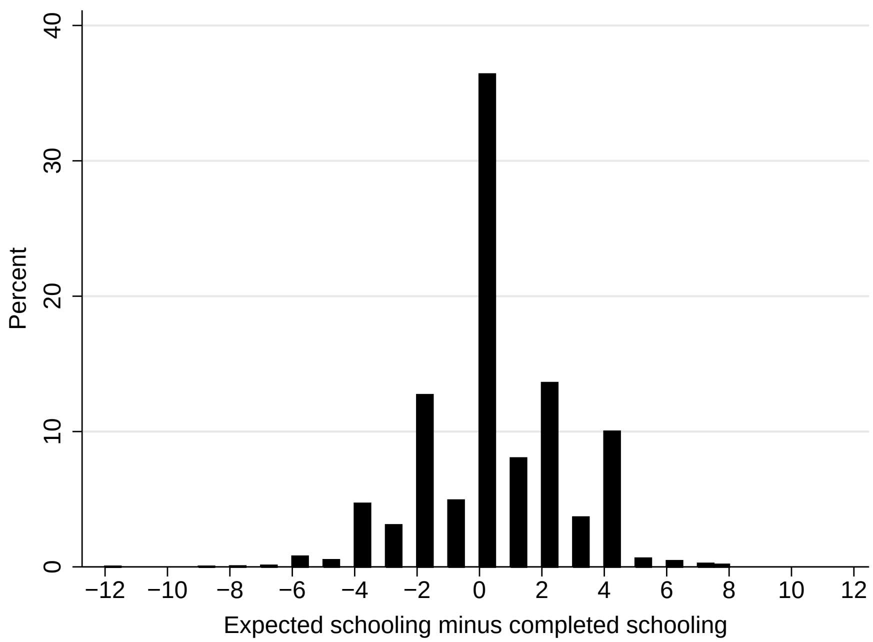
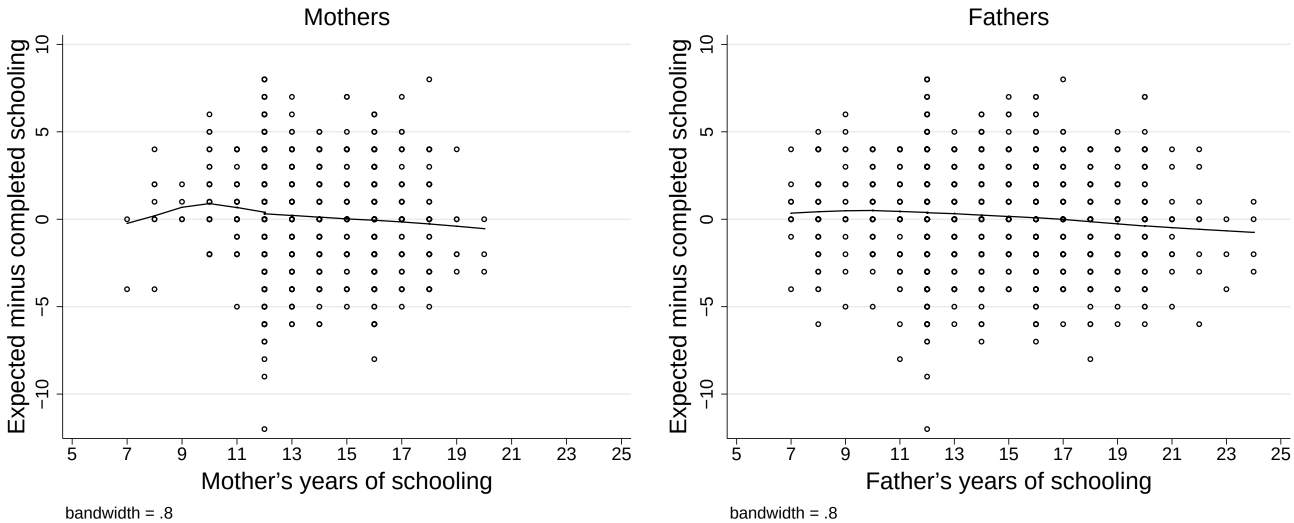


Figure 2.2: Lowess smooting: relation between parental schooling and prediction errors



Sample of selected children

So far, our findings with respect to the maximum likelihood and elimination approaches are based on the full sample of children while the results using the replacement method are generated using the smaller sample of one randomly selected child per family. It might be that the results reported in the previous sections are sensitive to the specific sample that is used. To check whether this is the case, Table 2.5 shows the same analyses as in Table 2.2 but now all results are based on the sample of selected children.

As can be seen in Table 2.5 the results are not very sensitive to the specific sample that is used to generate the results. Ignoring the censored observations give estimates that are downwardly biased. The maximum likelihood approach overestimates the effect of mother's and father's schooling, all the estimated biases are significantly different from zero. Also the elimination approach produces positively biased estimates, only for father's the differences are not significantly different from zero. The results whereby we replace the censored observations with parental expectations are the same as in Table 2.2 and show no significant bias.

Table 2.5: Estimates of the effect of parents' schooling on own birth children's schooling—sample of selected children

	Mobility Estimates without Censoring (WLS 2004)			Mobility Estimates with Censoring (WLS 1992)			Estimated Differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mother's	0.45***		0.24***	0.42***		0.22***	-0.03***		-0.02
schooling	(0.02)		(0.02)	(0.02)		(0.02)	(0.01)		(0.01)
Father's		0.35***	0.27***		0.32***	0.25***		-0.03***	-0.02***
schooling		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
N		4,097			4,097				
N^c		0			874				
CENSORED REGRESSION MODEL									
Mother's				0.53***		0.30***	0.08***		0.05***
schooling				(0.02)		(0.03)	(0.01)		(0.01)
Father's					0.39***	0.30***		0.04***	0.03***
schooling					(0.01)	(0.02)		(0.01)	(0.01)
N					4,097				
N^c					874				
EXCLUDING ALL CHILDREN YOUNGER THAN 25									
Mother's				0.50***		0.28***	0.05***		0.04***
schooling				(0.03)		(0.03)	(0.02)		(0.02)
Father's					0.36***	0.28***		0.01	0.01
schooling					(0.02)	(0.02)		(0.01)	(0.01)
N					2990				
N^c					167				
CENSORED OBSERVATIONS REPLACED WITH PARENTAL EXPECTATIONS									
Mother's				0.45***		0.23***	-0.00		-0.01
schooling				(0.02)		(0.02)	(0.01)		(0.01)
Father's					0.35***	0.27***		0.00	0.00
schooling					(0.01)	(0.01)		(0.01)	(0.01)
N					4,097				
N^c					874				

All regressions include additional controls or the child's age and gender. Standard errors (in parentheses) allow for correlation within families. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

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 mobility when none of the children has finished their schooling. Table 2.6 compares the mobility estimates on the uncensored sample with the results obtained when a researcher would have used the 1975 sample and replaced all observations by parental expectations.[11] For the sample of own birth children we find mobility estimates of 0.19 and 0.23 for mothers and fathers, respectively. For adoptees the effect of mothers schooling drops to 0.11 and is no longer significantly different from zero. The effect of father's schooling does not change much and remains statistically significant. Compared to the uncensored mobility results, these estimates 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.

Maximum likelihood and the elimination approach

The sensitivity analysis thus far 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 2.2 is not substantial.

We begin with the censored regression model. 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 10).

[11] The other two methods do not work with samples where all the observations are censored.

Table 2.6: Patience versus impatience

	Estimates without Censoring (WLS 2004)		Estimates with Expectations (WLS 1975)	
	(1)	(2)	(3)	(4)
Mother's schooling	0.24***	0.08	0.19***	0.11
	(0.02)	(0.06)	(0.02)	(0.16)
Father's schooling	0.26***	0.19***	0.23 ***	0.26**
	(0.01)	(0.04)	(0.01)	(0.10)
N	14,524	520	4,097	52
SAMPLES				
Own birth children	×	–	×	–
Adopted children	–	×	–	×

All regressions include additional controls for the child's age and gender.
Standard errors are in parentheses; * significant at 10% level,
** significant at 5% level, *** significant at 1% level.

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. The results show that the null hypothesis of homoskedasticity is not rejected but that the normality assumption is indeed rejected.[12]

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 up-

[12]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 0.366, the null hypothesis of homoskedasticity is therefore not rejected. The p-value of the skewness/kurtosis tests for normality is equal to 0.000, the null hypothesis of normality is therefore rejected

ward bias we observe in Table 2.2. 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.

2.6 Conclusion

Recent mobility studies that make a distinction between causation and selection 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.

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

Our basic result is that, net of assortative mating effects, positive parental schooling effects fall only little for fathers, but much more for mothers when we move from uncensored samples of own birth children to adoptees. This result appears to be fairly robust to the introduction of censored observations and the application of three correction methods. Parental schooling effects fall, but not by much, when mobility models are estimated on censored samples and rise, again not by much, when censored observations are tackled by either three correction methods. Of the three methods, the one that treats parental expectations as if they were realizations performs best. This result depends, however, on the degree of censoring. For samples that are largely incomplete the method does give a small (negative) bias.

Our results suggest that it doesn't 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. However, to draw more general conclusions from these remarkable findings we think that more research is needed on how parents form expectations and on how the replacement method works using different identification strategies on other data sets.