

NBER WORKING PAPER SERIES

ESTIMATING RETURNS TO SCHOOLING
WHEN SCHOOLING IS MISREPORTED

Thomas J. Kane
Cecilia Elena Rouse
Douglas Staiger

Working Paper 7235
<http://www.nber.org/papers/w7235>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1999

We thank Orley Ashenfelter, John Bound, Arie Kapteyn, Alan Krueger, Richard Murnane, and seminar participants at a number of universities and the NBER for helpful comments. All opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1999 by Thomas J. Kane, Cecilia Elena Rouse, and Douglas Staiger. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Estimating Returns to Schooling When Schooling is Misreported
Thomas J. Kane, Cecilia Elena Rouse, and Douglas Staiger
NBER Working Paper No. 7235
July 1999
JEL No. C1, I2, J31

ABSTRACT

We propose a general method of moments technique to identify measurement error in self-reported and transcript-reported schooling using differences in wages, test scores, and other covariates to discern the relative verity of each measure. We also explore the implications of such reporting errors for both OLS and IV estimates of the returns to schooling. The results cast a new light on two common findings in the extensive literature on the returns to schooling: “sheepskin effects” and the recent IV estimates, relying on “natural experiments” to identify the payoff to schooling. First, respondents tend to self-report degree attainment much more accurately than they report educational attainment not corresponding with degree attainment. For instance, we estimate that more than 90 percent of those with associate’s or bachelor’s degrees accurately report degree attainment, while only slightly over half of those with 1 or 2 years of college credits accurately report their educational attainment. As a result, OLS estimates tend to understate returns per year of schooling and overstate degree effects. Second, because the measurement error in educational attainment is non-classical, IV estimates also tend to be biased, although the magnitude of the bias depends upon the nature of the measurement error in the region of educational attainment affected by the instrument.

Thomas J. Kane
Kennedy School of Government
Harvard University
79 JFK St.
Cambridge, MA 02138
and NBER
tom_kane@harvard.edu

Cecilia Elena Rouse
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08544
and NBER
rouse@princeton.edu

Douglas Staiger
Department of Economics
Dartmouth College
Hanover, NH 03755
and NBER
douglas.o.staiger@dartmouth

I. Introduction

Over the last 40 years, a vast empirical literature has developed using survey data to estimate the relationship between wages and educational attainment. Until recently, the conventional wisdom held that ordinary-least-squares (OLS) estimates of the return to education were likely to be biased upward, due to a positive correlation between educational attainment and unobserved ability. In apparent support of the “ability bias” hypothesis, the early studies which controlled for ability-- either with test scores or with sibling and twin comparisons-- estimated somewhat lower returns to education.¹ More recently, a series of papers estimating the returns to education with instrumental variables (IV) have consistently found even higher returns than the original OLS estimates.² Measurement error has been cited as an explanation to reconcile these apparently conflicting results. Indeed, with classical measurement error in self-reported education, the OLS estimates with no ability measures could be biased either up or down, since the positive ability bias and negative measurement error bias could offset each other. With classical measurement error, the IV estimator using exogenous determinants of educational attainment would be consistent.

However, there is little reason to believe that the measurement error in self-reported education is classical. In fact, since the most widely used measures of educational attainment are categorical in nature (usually measured in discrete years or degrees), the measurement error generally will not satisfy the classical assumptions (Aigner, 1973). For example, one would expect the measurement error in any categorical measure to vary with the level of education reported, since individuals in the lowest education category can never under-report their education and those in the top education category cannot over-report. Unfortunately, without the classical assumptions traditional IV estimates of the return to education are no longer consistent. Moreover, one cannot place any *a priori* restrictions on the direction or magnitude of the OLS and IV bias except in special cases. Thus, the presence of measurement error in self-reports of educational attainment makes it, in principle, quite difficult to interpret many of the existing estimates of the returns to education. More generally, this problem arises whenever a categorical regressor is misreported. As a result, recent estimates of treatment effects using IV will be biased if there is misreporting of the treatment.

In this paper, we use a general method-of-moments (GMM) estimator to simultaneously estimate the returns to education and the distribution of reporting error in educational attainment. Our proposed method is appropriate when there are two categorical reports of educational attainment, with independent sources of potentially non-classical reporting error. Although we focus on measures of schooling, the estimator we propose is more generally applicable.

Our estimator has three distinct advantages over standard approaches. First, our estimator is consistent under weaker assumptions, and allows one to test directly some of the assumptions required by other estimators. For example, we do not need to place any strong restrictions on the form of the reporting error such as assuming classical measurement error (Ashenfelter and Krueger, 1994), assuming some form of symmetry in the measurement error (Card, 1996) or assuming that one measure is reported without error (Rodgers and Bruhl, 1997). Second, the estimator provides direct estimates of the measurement error in each measure, which may often be of independent interest. Finally, when other covariates are introduced, the GMM estimator is over-identified, and allows for a test of the key independence assumptions that are placed on the measurement error.

The next section outlines our empirical approach. Correcting for reporting error requires estimates of the distribution of the reporting error. We combine information on the number of people categorized differently by the two measures with information on the differences in their wages to estimate the distribution of reporting error in both measures and simultaneously produce “corrected” estimates of the value of schooling. The key assumption we make is that, conditional on a person’s actual schooling level, reporting errors are independent of wages and other covariates. Our estimation methodology is closely related to that developed by Jakubson (1986) for estimating union wage effects from panel data when union status is misreported and Black et. al. (1998) for examining measurement error in health insurance coverage. Like both Jakubson and Black et. al. we estimate the distribution of the reporting error directly, rather than from some external validation data set (e.g. see Card, 1996). However, our methodology is more general in that

we place no restrictions on the reporting error distribution, we consider the case of misreporting in a categorical variable with more than two categories, and we allow for covariates.

Sections III and IV of the paper apply our methodology to the educational attainment measures collected in two different datasets, the National Longitudinal Study of the High School Class of 1972 (NLS-72) and a subsample of high school drop-outs from the National Educational Longitudinal Study (NELS). We report a number of findings of substantive importance. First, self-reported measures appear quite accurate in reporting bachelor's degree attainment. We estimate that more than 95 percent of respondents with a bachelor's degree reported so accurately and less than 1 percent of those without a degree misreported having one. In fact, we estimate that self-reported measures of bachelor's degree completion were even more accurate than the transcript study which was conducted as part of the NLS-72 survey. Second, both self-reported and transcript measures are often inaccurate in reporting the quantity of schooling completed for those with between zero and four years of college completed. The data on years of schooling for those with no degrees are particularly suspect. For instance, we estimate that 6 percent of those with no college training self-report "some college" and that a similar proportion of those with "some college, no degree" misreport that they have no college. Similarly, we estimate that only slightly more than half (55 percent) of those with 1 year of college report their educational attainment accurately. Third, as a result of these errors in measurement for those with 1-3 years of college, OLS estimates seem to understate the earnings differentials associated with years of schooling completed. To the extent that degree completion is measured with less error than the years of college actually completed, our results suggest that "sheepskin effects"-- the estimated value of degree completion over and above the number of years of schooling attended-- may be overstated as well. For instance, we estimate that OLS estimates using self-reported schooling overstate the additional value of bachelor's and associate's degree (over and above the number of years of schooling completed) by roughly a factor of two. Fourth, based upon our results from the sample of high school drop-outs, self-reported measures also contain considerable measurement error for those completing less than 12 years of

schooling.

Finally, we find that reporting error in schooling leads to inconsistent IV estimates of the return to an additional year of school. For instance, if the measurement error in highest grade attended in the 1980 Census survey were similar to the measurement error we estimate for the NELS sample, IV estimates using quarter of birth as the instrument would overstate the return to education by roughly 34 percent. In contrast, IV estimates using years of education reported in college transcripts as the instrument for self-reported years of education understate the return to education by roughly 10 percent. Thus, at least in these cases, IV estimates of the return to education appear to be biased because of reporting error in schooling, but both the magnitude and direction of the bias depends upon the nature of the measurement error in the region of educational attainment affected by the instrument.

II. Estimation

The Empirical Model with a Binary Measure of Education and No Covariates

Our primary interest is in estimating the parameters in a regression of log wages on education when education is a categorical variable measured with error. To illustrate our approach, we consider a simple case in which educational attainment is measured as only two values (high and low) and there are no other variables entering the wage equation. In the next section, we derive a more general solution that allows for covariates and any number of educational categories.

Consider the following simple model³,

$$E(w|S^*) = \beta_0 + \beta_1 S^* \quad (1)$$

where w represents the logarithm of wages, and S^* is a dummy variable for the (unobserved) true level of schooling which is equal to 1 if education is "high" and 0 if education is "low." For simplicity we have suppressed subscripts (i) indicating that the variables are measured for individuals. Denote the mean of S^*

by μ . Since S^* is a binary variable, μ represents the probability of having "high" education, that is S^* equal to 1. In addition, we observe two measures of schooling S_1 and S_2 which are related to the true level of schooling S^* according to the following equations,

$$E(S_1|S^*, S_2, w) = \pi_{10} + \pi_{11}S^* , \quad (2)$$

and,

$$E(S_2|S^*, S_1, w) = \pi_{20} + \pi_{21}S^* . \quad (3)$$

Conditional upon the true value of schooling, S^* , we assume S_1 and S_2 are independent of each other and of w . In other words, given a person's true educational attainment, we assume that the measurement errors are independent of each other and of wages.

The assumptions embodied in equations (1)-(3) are the same as those made in models of classical measurement error, except that under classical measurement error one must also assume that $\pi_{11}=\pi_{21}=1$ and $\pi_{10}=\pi_{20}=0$. However, with categorical variables, the very existence of measurement error implies that $\pi_{11}<1$, $\pi_{21}<1$ and $\pi_{10}>0$, $\pi_{20}>0$. For example, let α_1 and α_2 be the error rates in S_1 (e.g. $\alpha_1 = \Pr(S_1=0|S^*=1)$, $\alpha_2 = \Pr(S_1=1|S^*=0)$). Then $\pi_{10}=\alpha_2$ and $\pi_{11}=1-(\alpha_1+\alpha_2)$, so that *any* non-zero error rates violate the classical assumptions. Therefore, the classical assumptions are not useful in the case of measurement error in a categorical variable and we must work under the more general assumptions embodied in equations (1)-(3).

Under these assumptions, common OLS and IV methods do not consistently estimate the returns to education. For example, some straightforward algebra shows that the OLS estimate of β_1 , in equation (1), converges to the following (Aigner, 1973):

$$\begin{aligned}\hat{\beta}_1^{\text{OLS}} &\Rightarrow \frac{\sigma_{w,S_1}}{\sigma_{S_1,S_1}} = \beta_1 [1 - Pr(S^*=1|S_1=0) - Pr(S^*=0|S_1=1)] \\ &= \beta_1 \left[1 - \frac{\alpha_1\mu}{\alpha_1\mu + (1-\alpha_2)(1-\mu)} - \frac{\alpha_2(1-\mu)}{\alpha_2(1-\mu) + (1-\alpha_1)\mu} \right],\end{aligned}\quad (4)$$

where the final expression is derived using Bayes law, and α_1 and α_2 are the error rates in S_1 as defined earlier. Since μ , α_1 , and α_2 are greater than zero and less than one by definition, the expression inside the square brackets is less than 1. Thus, OLS estimates of β_1 are biased toward zero, and cannot be corrected without knowledge of the error rates in reported education (α_1 and α_2) and the true probability that a person has “high” education (μ). In fact, if the measurement error is severe enough, the term in brackets in equation (4) can be negative and the OLS estimates wrong-signed.

More importantly-- particularly because it is often not recognized in the applied literature-- standard instrumental variables estimates of β_1 , which are consistent under classical measurement error, generally are not consistent when the measurement error is not classical.⁴ For example, the Two-stage-least-squares (2SLS) estimate of β_1 from regressing w on S_1 and using S_2 as an instrument converges to the following:

$$\hat{\beta}_1^{2\text{SLS}} \Rightarrow \frac{\sigma_{w,S_2}}{\sigma_{S_1,S_2}} = \beta_1 \frac{1}{\pi_{11}} = \beta_1 \frac{1}{1 - (\alpha_1 + \alpha_2)} . \quad (5)$$

Only in the case of classical measurement error (that is, when $\pi_{11}=1$) does 2SLS result in a consistent estimate of β_1 . In this case, however, $\pi_{11}<1$ when there is measurement error. As a result, the 2SLS

estimator will be upward biased (and in extreme cases of measurement error the 2SLS estimate will be wrong-signed.) Note that the bias of the 2SLS estimator depends only on the error in S_1 (the included regressor) and not on the measurement error in the instrument. In fact, in this example the bias of 2SLS is the same for *any* instrument (Z) satisfying the usual IV assumptions (i.e., Z is correlated with S^* but not with the error term in the outcome equation). Thus, the common use of 2SLS to instrument for a categorical variable results in biased estimates when the categorical variable is measured with error; and in this simple case of a binary regressor measured with error, the 2SLS estimates are upward biased (Card (1996) makes a similar point).

Of course, the inconsistency of the IV estimator derived above does not overturn the common textbook definition of the attributes of a valid instrument, that is, that a valid instrument must be correlated with the true value and uncorrelated with the measurement error. The problem is that any variable which is correlated with a categorical indicator of “true” schooling will generally also be correlated with the measurement error, since the measurement error itself is related to schooling. While this point may seem obvious, it is often overlooked (see, for example, the discussion in Fuller (1987)).

With a binary measure of education and no covariates, standard estimates will usually bound the true return to education (β_1): the OLS estimate is downward biased, and 2SLS estimate is upward biased (unless there is so much measurement error that the coefficients are wrong-signed). In the more general case, with multiple categories, one cannot generally sign the bias of OLS and 2SLS, so the standard estimates do not even provide bounds. The most common approach taken to obtain unbiased estimates uses external estimates of the measurement error probabilities (α_1 and α_2) to correct for the bias (see studies of union wage effects by Freeman (1984) and Card (1996)). These methods require either having a validation data set, in which one observes both the noisy measure and the actual measure, or placing other *a priori* restrictions on the form of the measurement error (Card, 1996 and Jakubson, 1986).

Our approach is to estimate the measurement error parameters directly from the data using a method of moments approach (see Jakubson (1986) and Black *et al.* (1998) for a closely related approach). The central idea behind this estimation strategy is to note that equations (1)-(3) depend on only seven parameters: β_0 , β_1 , π_{10} , π_{11} , π_{20} , π_{21} , and μ . At the same time, we have seven sample means that can be calculated in the data: The proportion of the sample in each cell from a 2×2 cross-tabulation of S_1 on S_2 (3 estimates, since the proportions sum to one); and the mean wage in each of these cells (4 estimates). The expected values of these seven sample moments are a function of the seven unknown parameters. Therefore, they provide a basis for a GMM estimator that is just-identified.

How does this estimator identify the measurement error? To see the intuition, suppose that one believes that S_2 has no error. If S_2 represents the “truth,” then we can estimate the error rates for S_1 from the cross-tab of S_1 and S_2 . Moreover, mean wages *among individuals with a given value of S_2* should be unrelated to S_1 . When this is not the case (e.g. if mean wages are higher when $S_1=1$ and $S_2=0$ compared to when $S_1=0$ and $S_2=0$), there is evidence that S_2 itself is measured with error. The observed mean wage in each cell is a weighted average of β_0 and $\beta_0 + \beta_1$ (the mean wages of those with true schooling equal to 0 and 1 respectively) where the weights are a function of the measurement error in each measure. By combining information from the cross-tabs with information on mean wages we are able to identify the extent of measurement error in the data.

The Case of Multiple Categories of Education and Covariates

The basic idea behind the GMM estimator for the simple case can be readily extended. Let S^* be a $1 \times J$ vector of dummy variables indicating which of J mutually exclusive and exhaustive categories is an individual’s true educational attainment (again, we suppress individual subscripts). For example, this could be a vector of dummies indicating whether an individual did not finish high school, was a high school graduate, attended college, or was a college graduate. For each level of schooling j , $S^*_j=1$ if the individual

has the particular level of schooling and $S_j^*=0$ if the individual does not.

We assume that the conditional expectation of log wages (w) is linear:

$$E(w|S^*,X) = S^*B_1 + XB_2 \quad (6)$$

where X represents a $1 \times K$ vector of additional covariates with mean zero, B_1 is a $J \times 1$ vector of parameters representing the return to each schooling level, and B_2 is a $K \times 1$ vector of parameters for the additional covariates.

Further, suppose that we have two $1 \times J$ “noisy” measures of S^* , S_1 and S_2 , such that the errors in S_1 and S_2 are independent of each other and other variables conditional on S^* (analogous to equations (2) and (3)). In this case,

$$E(S_1|S^*,X,S_2,w) = S^*\Pi_1 \quad (7)$$

and,

$$E(S_2|S^*,X,S_1,w) = S^*\Pi_2 \quad (8)$$

Equations (7) and (8) embody the key independence assumptions that lead to identification of the model. These equations state that the probability of reporting each level of schooling (for both measures of schooling) depends only on the true level of schooling.⁵ Both Π_1 and Π_2 are of dimension $J \times J$, and represent the error rates in each measure (i.e. $\Pi_1^{(ij)} = \Pr(S_1=j|S^*=i)$ where (ij) means the i th row and j th column of a matrix). Thus, the elements of any row of Π_1 or Π_2 sum to one.

Finally, we assume that the expectation for each true level of schooling (S^*) is linear in X :

$$E(S^*|X) = \delta_0 + X\delta_1$$

where δ_0 is $1 \times J$ and δ_1 is $K \times J$. Note that δ_0 represents the unconditional true proportion of individuals with

each schooling level (because X is mean zero) and is analogous to μ in our earlier example. Thus, the elements of δ_0 must sum to one. Similarly, since the elements of $E(S^*|X)$ must sum to one for any X , we impose the restriction that the elements of each row of δ_1 sum to zero.

Based on this model, we need to estimate $J+K$ parameters in the wage equation (β_1 and β_2), $J(J-1)$ parameters for each measurement error equation (Π_1 and Π_2), and $(J-1)(K+1)$ parameters that represent the correlation between the covariates and the true level of schooling (δ_0 and δ_1). Overall, there are $(2 \times J^2 - 1) + (K \times J)$ parameters to estimate.

We identify these parameters by fitting moments that are easily estimated from the data. As in the simple binary case with no covariates, we use the observed proportion of individuals with each combination of the two schooling measures, and the mean wages of individuals in each of these schooling-combination cells. In addition, we use the mean of the covariates (X) in each cell, and the covariance between the covariates and wages. As in the simple binary case, we infer the nature and level of the measurement error in schooling by observing how mean wages (and mean covariates as well) change when the two measures of schooling diverge.

Formally, let D be a $J^2 \times 1$ vector of dummy variables representing combinations of schooling levels from each measure of schooling. Suppose, for example, that the schooling levels range from 1 to J years of education. Therefore: $D_{11}=1$ if $S_1=1$ and $S_2=1$; $D_{12}=1$ if $S_1=1$ and $S_2=2$; and so forth. These dummy variables are mutually exclusive and exhaustive. We can write,

$$D = \begin{pmatrix} D_{11} \\ D_{12} \\ \vdots \\ D_{JJ} \\ D_{21} \\ \vdots \\ D_{JJ} \end{pmatrix} .$$

In addition, let T represent a $J^2 \times J$ matrix of the probabilities of observing each realization, conditional on the true level of schooling,

$$T = \begin{pmatrix} Pr(S_1=1, S_2=1|S^*=1) & \dots & Pr(S_1=1, S_2=1|S^*=J) \\ Pr(S_1=1, S_2=2|S^*=1) & \dots & Pr(S_1=1, S_2=2|S^*=J) \\ \vdots & \dots & \vdots \\ Pr(S_1=J, S_2=J|S^*=1) & \dots & Pr(S_1=J, S_2=J|S^*=J) \end{pmatrix}.$$

Because the measurement error for S_1 and S_2 are independent, we can write the elements of T in terms of Π_1 and Π_2 . For example, $Pr(S_1=1, S_2=1|S^*=1) = Pr(S_1=1|S^*=1)Pr(S_2=1|S^*=1) = \Pi_1^{(11)}\Pi_2^{(11)}$.

Using this notation we can compactly write the moments which allow us to identify the parameters of interest. Let a bar over a variable (or cross product) denote the sample average. Then the following moments hold:

$$E(\bar{D}) = T\delta'_0 \quad (12)$$

$$E(\bar{D}Y) = T(\delta'_0 + B_1) + T\delta'_1 \Sigma_{XX} B_2 \quad (13)$$

$$E(\bar{D}X) = T\delta'_1 \Sigma_{XX} \quad (14)$$

$$E(\bar{X}'Y) = \Sigma_{XX}(\delta_1 B_1 + B_2) \quad (15)$$

where Σ_{XX} is the variance-covariance matrix of X .

Equations (12)-(15) provide $(2J^2-1)+K(J^2+1)$ moments with which to identify our parameters. Thus,

the degree of over-identification is equal to $(2J^2-1)+K(J^2+1) - (2J^2-1)+(K \times J) = K(J^2 - J + 1)$. Without additional covariates ($k=0$), our estimates of the return to schooling and the measurement error are just-identified. However, with the addition of covariates the number of moments increases more rapidly than the number of parameters; as a result, the parameter estimates with covariates are over-identified. Note that with other covariates, we could still identify the nature of the measurement error in schooling, even if schooling does not have a causal effect on wages -- something which is not possible without covariates.

We implement an optimal minimum distance estimator to fit our parameters (Chamberlain, 1984). Note that with covariates, a rejection of the tests of over-identification provides evidence that the individual characteristics are correlated with the measurement error in schooling.

Although, as argued above, the parameters are identified without any additional restrictions, one can gain additional efficiency by using the prior knowledge that probabilities must be bounded by zero and one. Therefore, we used a logit specification for the measurement error matrices Π_0 and Π_1 such that,

$$P(S = i | S^* = j) = \frac{e^{\phi_j}}{1 + \sum_{l=1}^{k-1} e^{\phi_l}}.$$

A similar logit specification was used for estimating the true probability an individual is in each schooling level (δ_0). In the tables that follow, we report the implied probabilities (rather than the estimated logit parameters) with delta-method standard errors.

III. Results for High School Graduates from the Class of 1972

We first analyzed data from the National Longitudinal Study of the High School Class of 1972 (NLS-72), a longitudinal survey beginning with 22,652 seniors from the high school class of 1972. The NLS-72 is particularly well-suited to our purposes, given the availability of transcript data in the Post-secondary

Education Transcript Survey (PETS). In collecting the PETS data, transcript data were requested on all post-secondary schools reported by students during any of the 4 follow-up surveys between 1972 and 1979. Given the sample frame for the PETS survey, no post-secondary school entered by a youth after 1979 would have been included in the transcript survey. As a result, we focus on transcript-recorded and self-reported educational attainment as of 1979, when youth were already 7 years out of high school. In working with the transcript data, we ignored course-work or degrees earned for any term that began after 1979.⁶ If a student failed a course, the credit was not counted. Using a rule-of-thumb suggested by the American Council on Education, we counted quarter-system credits as two-thirds of a semester credit. When counting credits or finding degrees, we did not attempt to distinguish among different types of post-secondary institutions.

During the fourth follow-up survey, participants were asked “As of the first week of October, 1979, what was your highest level of college education?” They were given a choice among the following responses:

- o “This does not apply to me since I have not attended college.”
- o “Some, but less than two years of college.”
- o “Two or more years of college.”
- o “Finished college (four- or five-year degree).”
- o “Master’s degree or equivalent.”
- o “PhD or advanced professional degree.”

For much of our analysis of self-reported educational attainment, we relied upon responses to this question. However, we were also interested in associate’s degree completion, a category excluded from the above question. Two pages and eight questions later in the questionnaire, respondents were asked “Since high school, had you earned any certificate, license, diploma or degree of any kind prior to October 1979?” If they responded positively, they were asked to fill in the dates and fields for any of the following types of degrees: “certificate,” “license,” “2-year or 3-year vocational degree or diploma,” “2-year academic degree,” “4-year or 5-year college bachelor’s degree,” “master’s degree or equivalent” or other. We counted anyone

who reported a “2-year or 3-year vocational degree or diploma” or a “2-year academic degree” as having self-reported an associate’s degree.

When studying wages, we limited ourselves to those reporting wages between \$1.50 and \$80.00 per hour in 1986 (which, in 1986 dollars, was the 1st and 99th percentiles for those with positive earnings). For budgetary reasons, only a subsample of those individuals participating in prior waves were included in the 1986 sample frame. Moreover, certain subgroups -- Hispanics, teachers, persons self-reporting a 4-year college degree, and persons who were divorced, widowed or separated from their spouses -- were included with certainty. Because inclusion in the fifth follow-up was at least partially a function of self-reported educational attainment, we use the sampling weights to calculate the sample moments used in the GMM estimator. We also exclude those with missing transcripts except where otherwise noted.

Table 1 reports the proportion of the class of 1972 with bachelor’s degrees by 1979, according to both transcript-recorded and self-reported measures. The two measures were consistent in the vast majority of cases (97 percent of the sample was found along the diagonal). However, the transcript study did not always agree with what students reported. For instance, according to Table 1, 6 percent of those who reported having a bachelor’s degree did not have a bachelor’s degree on their transcript. If we were willing to take transcripts as representing the “truth,” we would be tempted to conclude that these 6 percent of respondents were not telling the truth.

However, Table 1 also reports the average log wages in 1986 for the same sample, by the level of self-reported and transcript-reported schooling. If the transcript measures represented the true level of schooling, we would expect to see no difference in earnings associated with self-reported schooling, holding transcript-reported schooling constant. Yet, among those for whom the transcript measure reported no bachelor’s degree, those who self-reported having a bachelor’s degree did have higher earnings than those who reported that they did not. Intuitively, as long as true schooling (not the variation in schooling due to measurement error) determines wages, such information ought to be useful in sorting out how much of the

Table 1
Sample Proportions and Mean Log Wages in 1986
by Self-Reported and Transcript-Reported Educational Attainment.

<i>Sample Proportions:</i> (Row Percentages) [Column Percentages]			
Self-Reported Schooling:			
PETS Schooling:	No Bachelor's Degree	Bachelor's Degree	Row Total:
No Bachelor's Degree	.725 (.977) [.992]	.017 (.023) [.062]	.742
Bachelor's Degree	.006 (.023) [.008]	.252 (.976) [.938]	.258
Column Total	.731	.269	1.000
<i>Mean Log Wages in 1986:</i>			
Self-Reported Schooling: (Standard Errors)			
PETS Schooling:	No Bachelor's Degree	Bachelor's Degree	Row Total:
No Bachelor's Degree	2.120 (.007)	2.309 (.037)	2.124 (.007)
Bachelor's Degree	2.280 (.063)	2.455 (.009)	2.451 (.009)
Column Total	2.121 (.007)	2.446 (.009)	2.209 (.006)

Note: Estimates are weighted, using 1986 weights. Educational attainment measured as of 1979. Average log hourly wages observed in 1986. The sample size was 5912.

Source: Authors' estimates using the NLS-72.

“truth” is reflected in each of the measures. The difference in mean wages between columns or between rows provides information about the relative verity of the two measures. This is the intuition behind the estimator we employ.

Log Wage Differentials for Self-reported and Transcript-reported Bachelor's Degree Attainment

Table 2 reports the results from the GMM estimator described above, as well as the OLS and 2SLS estimates one would obtain with self-reported and transcript-reported measures of bachelor's degree attainment. The 2SLS estimates were generated using the measure designated in the table as the regressor and the remaining measure as the instrument for the regressor. As reported in the first column of Table 2, the GMM estimate of the log wage differential for college graduates was .334. As we illustrated above, with a binary regressor measured with error, the OLS estimator will understate the difference in earnings associated with a bachelor's degree and the IV estimator will overstate that differential. Indeed, for both the transcript-reported and self-reported schooling measures, the OLS estimates (.326 and .324) fell below the GMM estimate, while the IV estimates (.348 and .341) overstated the wage differential associated with BA degree completion.

However, the magnitude of the difference between the GMM, OLS and IV estimates in Table 2 are quite small. Indeed, all four of the OLS and IV estimates fall within the 95 percent confidence interval surrounding the GMM estimate. The apparent reason, reflected in the GMM estimates of the measurement error, is that both the self-reported and transcript-reported educational attainment are estimated to be quite accurate for identifying bachelor's degree completion. For instance, we estimate that 99 percent of those without a bachelor's degree reported accurately in the NLS-72 survey. Moreover, 98.9 percent of those with a bachelor's degree are estimated to have accurately reported their educational attainment. Indeed, of the two measures, the transcript survey seemed to be less accurate, failing to capture 4 percent of those with a bachelor's degree.

Table 3
Estimates of the College Graduate Wage Differential with Covariates Included
(Dependent Variable: Ln Hourly Wage in 1986)

	Estimation Strategy				
	GMM	Transcript as Regressor		Self-Report as Regressor	
		OLS	TOLS	OLS	TOLS
Constant	2.143 (.007)	2.147 (.007)	2.142 (.007)	2.145 (.007)	2.140 (.008)
BA+	.248 (.016)	.241 (.016)	.262 (.017)	.241 (.015)	.258 (.017)
1972 Test Score	.076 (.009)	.080 (.008)	.075 (.008)	.078 (.008)	.074 (.008)
Female	-.328 (.012)	-.328 (.012)	-.327 (.012)	-.328 (.012)	-.327 (.012)
Black, Non-Hispanic	-.005 (.026)	-.008 (.025)	-.010 (.025)	-.013 (.025)	-.015 (.025)
Hispanic	.024 (.040)	.025 (.037)	.024 (.037)	.023 (.037)	.023 (.037)
Other, Non-Hispanic	.016 (.031)	.011 (.029)	.010 (.029)	.012 (.029)	.012 (.029)
Estimated Error Probabilities for the Transcript and Self-Reported Schooling Levels Using GMM					
Conditional upon Actual Schooling Level:					
		No BA		BA+	
Transcript	No BA	.998 (.000)	.050 (.007)		
	BA+	.002 (.000)	.950 (.007)		
Self-Reported	No BA	.997 (.002)	.015 (.003)		
	BA+	.003 (.002)	.985 (.003)		
Estimated Proportion in Each Category		.731 (.005)	.269 (.005)		

Note: Standard errors are in parentheses. There were 5912 observations. The p-value on the test of over-identifying restrictions was .731. The test score was standardized to have mean zero and variance one.
Source: Authors' calculations using the NLS-72.

The results in Tables 1 and 2 illustrate another important assumption underlying our approach to identifying measurement error. We assume that the measurement error in the two measures of educational attainment are independent. This assumption, along with the fact that the two measures took on the same values for 97.7 percent of the sample, together implied very little measurement error in either measure. If the measurement error had been large for either measure, and if the measurement errors for the two measures were independent, we would have expected to see a larger proportion of the sample off of the diagonal in Table 1. The fact that we did not led us to infer that there was very little measurement error in the data.

Table 3 reports similar results for the same sample with covariates, including indicators for gender and race/ethnicity, as well as each individual's score on a standardized test (including 4 equally weighted subsections of vocabulary, reading, letter groups and math) from the base year. The inclusion of covariates has little effect on the estimated measurement error probabilities for each measure of schooling. Not surprisingly, controlling for these covariates reduces the estimated return to a BA. To the extent that these covariates are more strongly related to "true" schooling than to the measurement error, the inclusion of covariates typically amplifies the bias due to errors-in-variables, since a larger portion of the orthogonal variation in educational attainment will be made up of measurement error when covariates are included than when they are excluded. However, the bias in the OLS estimates of BA degree completion -- .241 versus a GMM estimate of .248-- is still quite small with covariates.

Over-identification Tests

The fundamental assumption underlying the GMM estimator is that only "true" educational attainment should be related to mean wages-- i.e. that misstatements should be unrelated to wages. As discussed briefly in the theory section above, we have also assumed that the measurement error is independent of each covariate (although we allow each of the covariates to be linearly related to true educational attainment). Clearly, this is a rather strong assumption. For instance, since gender is a

covariate, we are assuming that men are not more likely to misreport their education than women. Fortunately, this is also a testable assumption, since the GMM estimator is over-identified when covariates are included. Intuitively, we could construct Table 2 using each of the covariates (in addition to log wages) as the dependent variable. If we would have gotten different estimates of the measurement error components using the covariates as the outcome rather than wages, the test of over-identifying restrictions reported at the bottom of Table 3 should fail. In fact, the p-value on the test was .731 in Table 3, meaning that we could not reject the hypothesis that the measurement error is uncorrelated with each of the covariates.

To evaluate the power of the over-identification test, we randomly eliminated varying proportions of the bachelor's degrees for men, re-estimated the model and then re-calculated the test of over-identifying restrictions. Since such simulations include a "manufactured" correlation between the measurement error and one of the covariates (gender), the test of over-identifying restrictions should fail. We randomly eliminated 1 percent, 5 percent and 10 percent of the bachelor's degree reported on the transcripts by men in the sample. While the test would not have discovered a 1 percent difference in the error rate for men and women, even a 5 percent divergence from the independence assumption for a single measure was large enough to lead us to fail the over-identification test. Put another way, the overidentification test was sensitive enough to discern a difference in misreporting by gender that affected only 48 observations in a sample of 5912.

Estimates for the Subsamples with Missing Transcripts

As described in the data section, the above estimates are limited to the sample with no missing transcripts in the PETS follow-up. Therefore, as a second check on the ability of our approach to detect measurement error, we considered the effect of including individuals for whom we know that the transcript information is incomplete. First, we estimated the GMM model for the subsample of respondents who were missing at least one postsecondary transcript because a school replied that the individual had never attended.

Solely for the sake of mnemonics, we call this the “liar” subsample (although as we will see, the label may not be deserved). Second, we used the subsample who were missing postsecondary transcripts for any other reason, such as because the school had closed, had lost their records, or simply did not reply.

The results are in Table 4. The estimator correctly recognized that something was amiss in the transcript measures. For instance, in the subsample with transcripts missing due to school non-response, the GMM estimator estimated that 26.3 percent of those with a BA had the BA missing from the transcript data. It continued to find little evidence of error in the self-reported data. Interestingly, even in the “liar” subsample-- those for whom at least one college reported that the person never attended-- the GMM estimator implied that it was primarily the transcript measure that was in error and not the self-reported measure. Indeed, based upon the wages of those who claimed to have a bachelor’s degree, the estimator apparently concluded that the schools, and not the students, had been mistaken. Among those whose names the schools claimed not to have in their records, one-sixth (15.8 percent) of those with true BA’s are estimated to have been missing a transcript-reported BA. Moreover, despite the error in the measures, the GMM estimates of the value of a BA degree for those with missing transcripts were not significantly different from the full-sample estimate of .248 in Table 3.

Results including Some College and AA degree Completion

Table 5 reports the log wage differentials using 4 categories of educational attainment: bachelor’s degree recipients, associate’s degree recipients, those with some college (no associate’s degree) and those with no college. Again, bachelor’s degree attainment seems to be well-reported in the self-reported data. The confusion seems to be in drawing the line elsewhere-- between those with no college, some college or an associate’s degree. According to the estimates in the bottom of Table 5, 6.4 percent of those with no college self-report some college. Also striking is the apparent failure of the transcript data to successfully identify those with associate’s degrees. The estimates in Table 5 suggests that 34 percent of those with

Table 4
Estimated College Wage Differentials for Samples with Missing Transcripts

	“Liar” Subsample	“Other Missing Transcript” Subsample
	GMM	GMM
Constant	2.164 (0.008)	2.164 (0.051)
BA+	0.156 (0.060)	0.225 (0.147)
Covariates Included?	Yes	Yes
p-value of test of Over-Identification	.999	.845
N:	655	662

Estimated Error Probabilities for the PETS and Self-Reported Schooling Levels Using GMM

Reporting:		“Liar” Subsample Actual Schooling		“Other Missing” Subsample Actual Schooling	
		No BA	BA+	No BA	BA+
PETS	No BA	1.010 (0.067)	0.158 (0.030)	1.010 (0.072)	0.263 (0.036)
	BA+	-0.010 (0.067)	0.842 (0.030)	-0.010 (0.072)	0.737 (0.036)
Self- Report	No BA	0.977 (0.017)	0.019 (0.094)	0.968 (0.022)	0.002 (0.153)
	BA+	0.023 (0.017)	.981 (0.094)	0.032 (0.022)	0.998 (0.153)

Note: Standard errors are in parentheses. Covariates are the same as in Table 3. The p-value on the test of over-identifying restrictions was .999 for the “liar” subsample and .845 for the “other missing transcript” sample.

Source: Authors’ calculations using the NLS-72.

Table 5. Estimates of the Wage Differentials for College Graduates and those with Some College with Covariates Included (Dependent Variable: Ln Hourly Wage in 1986)

		Estimation Strategy			
		Transcript as Regressor		Self-Report as Regressor	
GMM		OLS	TOLS	OLS	TOLS
Some College, No AA	.123 (.027)	.121 (.015)	.141 (.023)	.106 (.016)	.167 (.024)
Completed AA Degree	.216 (.034)	.133 (.029)	.278 (.042)	.185 (.024)	.156 (.034)
BA+	.330 (.027)	.313 (.018)	.353 (.020)	.315 (.018)	.353 (.020)
Estimated Error Probabilities for the Transcript and Self-Reported Schooling Levels Using GMM					
Conditional upon Actual Schooling Level:					
		No College	Some Coll, No AA	AA degree	BA+
Transcript	No College	.933 (.024)	.070 (.014)	.027 (.017)	.012 (.003)
	Some College, No AA	.066 (.009)	.920 (.018)	.339 (.069)	.027 (.006)
	AA Degree	.000 (.001)	.000 (.010)	.631 (.074)	.004 (.003)
	BA+	.000 (.024)	.010 (.006)	.003 (.005)	.957 (.008)
	Self-Reported	No College	.924 (.009)	.061 (.015)	.000 (.003)
	Some College, No AA	.064 (.009)	.928 (.033)	.075 (.045)	.008 (.003)
	AA Degree	.012 (.003)	.000 (.031)	.915 (.047)	.002 (.002)
	BA+	.000 (.001)	.011 (.007)	.010 (.014)	.990 (.026)
Estimated Proportion in Each Category		.419 (.011)	.239 (.012)	.073 (.010)	.269 (.009)

Note: Standard errors in parentheses. There were 5912 observations. Same covariates as in Table 3 were included. The p-value on the test of over-identifying restrictions was .312.

Source: Authors' calculations using the NLS-72.

associate's degrees are mis-categorized as "some college, no degree" in the transcript data. This divergence is reflected in the OLS and IV estimates of the value of an associate's degree. The GMM estimate of the log wage differential between associate's degree holders and those with no postsecondary education is .216. However, the OLS and 2SLS estimates using the transcript measure as the regressor would have been .133 and .278 -- 40 percent too low in first case and 29 percent too high in the second. The OLS and 2SLS estimates of the wage differential associated with attending "some college" are similarly misstated. For instance, using self-reported education as a regressor, we would have alternatively under-reported the estimate by 15 percent using OLS (.106 versus .123), and over-reported by 34 percent using 2SLS (.167 versus .123).

Impacts of Postsecondary Education on 1979 Test Scores

If respondents misreport degree completion to their employers as well as to survey researchers, and some proportion succeed in the deception, misreported BA's may indeed be related to market wages, thereby violating one of our assumptions for identifying the errors in both transcript-reported and self-reported schooling. Recall that we rely heavily on the differences in wages for those with differences on the transcript and self-reported schooling to infer the extent of error in each measure. Given our assumption that only the truth would matter for wages, our method would give the benefit of the doubt to the self-reported measure and wrongly conclude that the transcript study missed some legitimate bachelor's degree recipients.

However, while it is plausible that such deception would have lingering effects on wages, it is less plausible that misreported schooling would be related to test score gains. Fortunately, a subsample of respondents were given a second test of math and reading skills at the time of the 1979 follow-up, using a subset of the items included in the base-year test. Therefore, in addition to using wages as the outcome variable, we use the change in test scores between 1972 and 1979 to identify the measurement error in the transcript and self-reported schooling measures. (We have standardized the change to have a mean of zero

and a variance of one.)

Table 6 repeats the exercise in Table 5, using gains in test scores between 1972 and 1979 as the dependent variable. Interestingly, the nature of the measurement error is quite similar using either wages or gains in test scores between 1972 and 1979 to weigh the plausibility of transcript and self-reported data. For instance, in the self-reported educational attainment, 6 to 10 percent of those with no college mistakenly report “some college.” However, the self-reported data on associate’s degree attainment is estimated to be remarkably accurate, with between 87 and 92 percent of those with associate’s degrees correctly reporting their educational attainment. In contrast, the results suggest that the transcript data are particularly poor in identifying associate’s degree completion, missing between 34 and 21 percent of those with such degrees when wages and test score gains are used to identify misreporting.

Estimating Returns per Year of College Completed and Degree Effects

For roughly 5 decades between 1940 and 1990, the Bureau of the Census measured educational attainment by asking respondents to report years of school attended and whether or not they had completed the highest grade attended. As a result, much of what we know about the wage differentials associated with educational attainment is based on such data. In this section, we attempt to learn about the error properties of such measures.

Unfortunately, the NLS-72 does not inquire about educational attainment in precisely the same way as the Census bureau. Rather than being asked about the highest grade attended and whether or not they had completed that grade, respondents in the 1979 wave of the survey were asked, “Since leaving high school, about how many credits which can be used for a 4-year college Bachelor’s degree had you earned by October 1979?” In responding, respondents were given the choice of reporting “Number of quarter hours,” “Number of semester hours” and “Number of other type of credits.” Respondents were allowed to use more than one line when applicable. Later in the survey, respondents were asked to report credits completed toward a

Table 6. Estimates of the Test Score Differentials for College Graduates and those with Some College with Covariates Included (Dependent Variable: Standardized Change in Test Scores 1979-1972)

		Estimation Strategy			
		Transcript as Regressor		Self-Report as Regressor	
GMM		OLS	TOLS	OLS	TOLS
Some College, No AA	.213 (.052)	.151 (.046)	.241 (.071)	.165 (.047)	.221 (.074)
Completed AA Degree	.096 (.097)	.106 (.105)	.210 (.166)	.175 (.083)	.128 (.127)
BA+	.202 (.050)	.192 (.054)	.217 (.060)	.190 (.053)	.229 (.060)
Estimated Error Probabilities for the Transcript and Self-Reported Schooling Levels Using GMM					
Conditional upon Actual Schooling Level:					
		No College	Some Coll, No AA	AA degree	BA+
Transcript	No College	.903 (.036)	.075 (.060)	.000 (.000)	.000 (.029)
	• Some College, No AA	.095 (.033)	.925 (.089)	.212 (.118)	.055 (.039)
	AA Degree	.000 (.006)	.000 (.000)	.788 (.128)	.004 (.004)
	BA+	.001 (.012)	.000 (.071)	.000 (.065)	.941 (.051)
Self-Reported	No College	.895 (.039)	.000 (.062)	.023 (.068)	.000 (.024)
	Some College, No AA	.097 (.034)	.922 (.069)	.111 (.034)	.019 (.088)
	AA Degree	.007 (.004)	.061 (.024)	.866 (.074)	.011 (.015)
	BA+	.002 (.014)	.016 (.032)	.000 (.000)	.970 (.100)
Estimated Proportion in Each Category		.478 (.031)	.268 (.037)	.048 (.009)	.205 (.025)

Note: Standard errors in parentheses. There were 2311 observations. Same covariates as in Table 3 were included. The p-value on the test of over-identifying restrictions was .983.

Source: Authors' calculations using the NLS-72.

graduate degree using a very similar format.

We tabulated student responses to these questions, counting a “quarter” system credit hour as two-thirds of a semester credit hour. We also compared student responses to what was available in the transcript survey, using the same criterion to convert quarter-system credits to semester credits. (We limited both samples to those with no missing transcripts.) We divided the sample into one of 7 mutually exclusive categories: 0, 1, 2, 3 or 4+ years of college with no degree, Associate’s degree being highest reported degree and Bachelor’s degree being the highest degree.⁷ Based on these education categories we estimated two specifications. In the first specification, wages depended linearly on years of education and other covariates, with years of education set equal to years of college for those without degrees, 2 for those with associate’s degrees and 4 for those with bachelors degrees. In the second specification, we also included indicators for those with associate’s (AA) and bachelor’s (BA) degrees. The coefficients on the AA and BA dummies are typically referred to as “sheepskin effects” – the difference in earnings associated with having completed a degree over and above the earnings one would expect based on the number of years of schooling completed.

The results are reported in Table 7. Given the large number of categories, the measurement error matrices are reported in Appendix Table 1. The top panel of Table 7 reports estimates from the linear specification (without sheepskin effects), while the bottom panel reports the results with sheepskin effects included. In the specifications without sheepskin effects, the GMM estimate of the return to a year of education is higher than both the OLS and the IV estimates, with OLS estimates biased downward about 20 percent, and IV estimates biased downward by 5-10 percent.⁸

In the specifications in the bottom of Table 7, that allowed for sheepskin effects, the GMM estimator of the differential per year of college completed (.077) is 71 percent larger than the OLS estimate based upon transcript measures of schooling and 43 percent larger than the OLS estimate based upon self-reported schooling. The reason, apparently, is the large amount of measurement error among those with 0, 1, 2 and 3 years of schooling. Meanwhile, the GMM estimates of the associate’s and bachelor’s degree effects (that

Table 7
Estimates of the Wage Differentials by Year of College Completed and Degree Completion
(Dependent Variable: Ln Hourly Wage in 1986)

	GMM	Transcript as Regressor		Self-Report as Regressor	
		OLS	2SLS	OLS	2SLS
<i>A. Linear Specification</i>					
Years of College	.086 (.025)	.069 (.004)	.082 (.005)	.071 (.004)	.079 (.005)
<i>B. Linear Specification with Degree Effects</i>					
Years of College	.077 (.024)	.045 (.008)	.067 (.014)	.054 (.008)	.057 (.013)
AA Degree	.040 (.047)	.029 (.033)	.133 (.058)	.072 (.027)	.019 (.047)
BA Degree	.034 (.063)	.111 (.031)	.061 (.052)	.076 (.030)	.099 (.051)

Note: Standard errors in parentheses. There were 7 educational attainment categories: 0, 1 yr (no degrees), 2 years (no degrees), 3 years (no degrees), 4 or more years (no degrees), Associate's degree, Bachelor's degree. In estimating wage differentials, years of college were set equal to 2 and 4 for those reporting associate's and bachelor's degrees respectively. In other words, the coefficient on the degree effects represent the differences in earnings from those with 2 or 4+ years of college and no degree. There were 5281 observations. The same covariates as in Table 3 were included. The p-value on the test of over-identifying restrictions was .953 and .952 for specification A and B. Measurement error probabilities for specification B are reported in Appendix Table 1.

Source: Authors' calculations using the NLS-72.

is, the differences in earnings for those with associate's and bachelor's degrees relative to those with 2 and 4 years of schooling with no degrees) are about half the size of the OLS-estimated degree effects using self-reported schooling. In other words, the commonly observed finding of "sheepskin effects" in OLS returns to schooling-- differences in earnings between those reporting degrees and those with similar amounts of schooling without degrees-- is at least partially due to the nature of the reporting error in schooling. Because people report degrees more accurately than they do years of schooling completed when they do not have degrees, OLS estimates tend to overstate degree effects and understate earnings differentials per year of schooling completed.

V. Estimating Measurement Error in Reported Schooling for a Sample of High School Drop-outs

The evidence thus far suggests that OLS and 2SLS estimates of the returns to schooling are potentially misleading for estimates that are identified from levels of education for which individuals are likely to misreport their attainment; i.e., non-degree years. Another group for whom the measurement error in schooling may be quite high are high school dropouts. If so, this could have important implications for some recent IV estimates of returns to schooling. For example, in an influential paper published in 1991, Angrist and Krueger exploit the fact that individuals born in the first quarter of the year complete less schooling than individuals born later in the year (presumably because of the interaction between regulations governing minimum age of school entry and compulsory schooling laws) in order to estimate the payoff to schooling. They argue that since most school districts require those born on January 1 to wait and enter school one year later than those born on December 31, those born in the first quarter of the year reach the age of compulsory schooling after having completed 1 year less schooling than youth born at the end of the previous year. Assuming that those born in the first quarter are otherwise similar to those born in the last quarter of the year, Angrist and Krueger used such fortuitous timing as a "natural experiment" with which to identify the impact of such compulsory schooling on the earnings of men born between 1920 and 1950.

Their estimates are identified from those who would have dropped-out of high school were it not for the compulsory schooling laws.

Unfortunately, the NLS-72 data do not allow us to test the measurement error in educational attainment for those affected by compulsory schooling laws, simply because the survey began with a sample of high school seniors -- already beyond the age of compulsory schooling. Since compulsory schooling laws usually require youth to remain in school until age 16 or 17, the quarter-of-birth "experiment" is usually only relevant for those who would be dropping out of school between the 8th and 12th grades.

To learn something about the nature of error in self-reported educational attainment for those completing less than a high school degree, we employ the NELS which began with a sample of 8th grade students in 1988. During the spring term of the 1991-92 academic year, any student who was not enrolled in a high-school program and did not already hold a high school degree was asked to report their grade level at the time of dropout. In the following months, the students' parents were also surveyed and asked to report the students' grade level at dropout.⁹ In this section, we use students' and parents' reports of highest grade level attended to learn something about the nature of measurement error in educational attainment for those completing less than 12 years of school.¹⁰

Of the 1310 drop-outs for whom both students and parents reported highest grade attended, students' and parents' reported the same educational attainment for only slightly over half the sample (57 percent). One-third of the sample (35 percent) disagreed by 1 year and one-twelfth (8 percent) differed by 2 or more years. However, there were 2 empty cells in the cross-tabulation of student reported and parent-reported schooling: among those where the student reported having attended 12th grade, no parents reported students having attended 8th or 9th grade. In order to ensure identification, we randomly assigned 5 sample members into each of those two cells.¹¹

We employed the same GMM estimator to generate estimates of the measurement error in educational attainment for the drop-outs in the NELS sample, using grade-point-average in the 8th grade as

the dependent variable used to calculate the relative verity of the two measures. The estimates of measurement error are reported in Table 8. Although youth's reports are estimated to have been more accurate than their parents in 4 out of 5 categories, the estimator implies a considerable amount of measurement error for both students and parents, with diagonal elements of the measurement error matrix ranging from .54 to .98.

What would such estimates imply about IV estimates using quarter-of-birth as an instrumental variable? Table 9 reports the results of using quarter of birth to instrument for years of schooling attended for men born between 1930 and 1939 in the 1980 census who reported attending between 8 and 12 years of schooling.¹² Relative to those born in the 2nd through 4th quarters, those born in the first quarter had annual earnings .0045 log points lower and reported attending .0353 fewer years of schooling. These two estimates would have implied a return per year of schooling of .127-- quite similar to the estimate reported in Angrist and Krueger (1991), albeit with a larger standard error due to our restriction limiting the sample to those who attended between 8 and 12 years of schooling.

The bottom panel of Table 9 reports differences at each grade level in highest grade attended by quarter of birth. For instance, those born in the first quarter were .6 percentage points more likely to report having completed only the 8th grade and 1 percentage point less likely to have reported completing the 12th grade. Since these observed differences in educational attainment at each grade level are simply the product of the measurement error matrix (Π) and the actual change at each level, we can estimate the "actual" change in educational attainment at each grade level by pre-multiplying the vector of observed impacts for each year of schooling attended by the inverse of the measurement error matrix (Π^{-1}).

Using the estimates of measurement error in educational attainment reported by the students themselves, we use this identity to estimate the "actual" differences in educational attainment by quarter of birth at each grade level in the bottom panel of Table 9. For instance, our estimates suggest that the "true" impacts of quarter of birth were more concentrated at grade 8 and grade 12 than the "observed" impacts.

Table 8
Measurement Error in Highest Grade Attended
Among High School Dropouts

		Conditional upon Schooling Level at Dropout:				
		Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Student	Grade 8	.625 (.235)	.055 (.044)	.000 (.027)	.000 (.000)	.000 (.003)
	Grade 9	.247 (.212)	.779 (.117)	.098 (.259)	.000 (.000)	.000 (.033)
	Grade 10	.000 (.094)	.166 (.117)	.546 (.181)	.062 (.118)	.069 (.053)
	Grade 11	.096 (.070)	.000 (.000)	.326 (.136)	.938 (.224)	.107 (.133)
	Grade 12	.033 (.081)	.000 (.020)	.030 (.025)	.000 (.169)	.825 (.129)
Parent	Grade 8	.619 (.189)	.084 (.092)	.000 (.000)	.000 (.007)	.016 (.012)
	Grade 9	.135 (.189)	.732 (.593)	.102 (.069)	.010 (.045)	.013 (.010)
	Grade 10	.146 (.096)	.166 (.419)	.634 (.095)	.000 (.159)	.037 (.027)
	Grade 11	.035 (.067)	.000 (.216)	.225 (.092)	.892 (.213)	.273 (.110)
	Grade 12	.065 (.051)	.018 (.062)	.038 (.042)	.098 (.116)	.661 (.104)
Estimated Proportion at Each Grade Level:		.035 (.021)	.159 (.153)	.373 (.186)	.200 (.097)	.233 (.059)

Note: Standard errors in parentheses. There were 1310 observations. The dependent variable was student-reported grades in 8th grade. There were no covariates.

Source: Authors' calculations using the NELS.

Table 9

Inferring IV Bias in the Census using Estimated Measurement Error from the NELS

	Reduced Form:		Yrs of Schooling Attended	Second Stage:	
	Ln of Annual Earnings (OLS)	Years of Schooling Attended (OLS)		Ln of Annual Earnings (OLS)	Ln of Annual Earnings (IV)
Born 1st Quarter	-.0045 (.0035)	-.0353 (.0072)		.075 (.001)	.127 (.098)

	Reported Highest Grade Attended by Qtr of Birth		Observed Difference: (ΠA)	Estimated Actual Difference ($\hat{\Pi}^{-1}\Pi A$)
	Born 2-4th Quarter:	Born 1st Quarter:		
Grade 8	.0833	.0895	.0062	.1010
Grade 9	.0610	.0626	.0016	-.0021
Grade 10	.0966	.0992	.0026	.0073
Grade 11	.0782	.0784	.0002	-.0018
Grade 12	.6809	.6703	-.0106	-.0135
Avg. Years of Schooling:	11.212	11.177	-.035	-.047

Note: Based on a sample of 178,837 men in the 1980 Decennial Census, born between 1930 and 1939, whose highest grade attended was between 8 and 12 years of schooling. Standard errors are in parentheses.

Instead of being .6 percentage points more likely to have completed 8th grade, those born in the first quarter are estimated to have been 1 percentage point more likely. The reporting error effectively smoothed much of the impact over the intervening grades 9, 10, 11. With the type of measurement error reported in Table 8, the observed impacts of quarter of birth on years of schooling completed would have been 34 percent too small-- .035 years as opposed to .047 years.

Recall that the simple IV estimate of the return to schooling can be estimated as the ratio of wage differences corresponding to quarter of birth to the differences in average educational attainment by quarter of birth. While measurement error in educational attainment does not affect the numerator of this ratio, the denominator is understated in this case. As a result, the IV estimate would be estimated to be roughly 34 percent overstated, due to the non-classical nature of the measurement error in reported educational attainment among high school drop-outs. In other words, the IV estimate of the return to schooling based on the quarter of birth-- adjusted for the type of non-classical measurement error we observed in the NELS sample-- would have been .095, 25 percent higher than the OLS estimate but 34 percent lower than the IV estimate.

Given the difference in the timing of the surveys (the 1980 census captures those who were adults much later in life, while the NELS survey captured measurement error among recent high school drop-outs), our estimates are not directly comparable to the results in Angrist and Krueger (1991). Moreover, many of the individual components of the measurement error matrix in Table 8 are rather imprecise, further suggesting a note of caution in extending these results. However, the calculation above is useful in two respects: First, it illustrates the fact that even a modest degree of non-classical measurement error can have very large impacts on IV estimates. Second, both the magnitude and direction of the bias (when there are more than 2 categories) depends crucially on where in the range of educational attainment a particular intervention has its effects. Even if educational attainment includes a broad range of categories-- say grade 8 through grade 16-- it is only the measurement error in education at the margin along which a

particular intervention has its impact that matters, not the measurement error at all levels of educational attainment.

VI. Conclusion

Employing the assumption that only "true" schooling, and not measurement error, would be related to wages or test scores (or other covariates), we infer the extent of measurement error in self-reported and transcript-based measures of educational attainment. Perhaps surprisingly, our results suggest that transcript studies are often subject to at least as much measurement error as survey questions. Indeed, survey questions can be even more accurate than transcripts for capturing certain facts, such as bachelor's degree attainment.

The proposed methodological approach sheds new light on two common findings in the vast literature on educational attainment: "sheepskin" effects and large IV estimates of the returns to schooling. While self-reported schooling measures are fairly accurate in determining who has not attended college and who has completed their bachelor's degree, they are much less accurate in distinguishing among those with 1, 2 and 3 years of college. To the extent that degree attainment (or having 0 or 4 or more years of college) is simply more accurately measured than whether one has completed 1, 2 or 3 years of college, OLS estimates are likely to overstate the marginal benefit of completing a degree. Unfortunately, the question regarding the number of years completed in college on the NLS-72 is not directly comparable to the widely-used *Current Population Survey* question on educational attainment. As a result, we cannot test the validity of that question directly. However, to the extent that traditional educational attainment questions have had more difficulty discerning among those with 1, 2 and 3 years of college than identifying those with 0 and 4 years of college, OLS estimates based on such measures are likely to understate the value of a year of education and overstate "sheepskin" effects.

A number of recent papers have reported IV estimates of the return to education that are larger than

OLS estimates. For example, Angrist and Krueger (1991) use the interaction between quarter-of-birth and compulsory schooling laws to identify the payoff to schooling, while Kane and Rouse (1993) and Card (1993) use the distance of one's high school from the nearest two-year or four-year college as an instrument for schooling. This literature typically portrays instrumental variables as the preferred estimator in the face of measurement error. Our results suggest that *both* the IV and OLS estimates may be biased by measurement error in educational attainment, since such errors are typically not classical in nature.

The problem of measurement error may be particularly acute precisely in the range of educational attainment where most of these "natural experiments" would have had their greatest impact. For instance, compulsory schooling laws are likely only to be binding for those dropping out between 9th and 10th or 10th and 11th grade. Likewise, living close to a two-year or four-year college may raise the likelihood of completing 1 year of college-- if it has any effect at all-- but would have little impact on bachelor's degree attainment. While degree attainment seems to be measured quite accurately in self-reported measures, neither instrument would be expected to have much of an effect on the proportion of youth finishing high school or college degrees. Rather, because they operate on more nebulously defined margins of educational attainment-- such as 1, 2 or 3 years of high school or 0 or 1 year of college-- such estimates are likely to be particularly problematic. In the presence of measurement error, the range over which such "natural experiments" have their impacts is fundamentally important.

References

- Aigner, Dennis J. "Regression with a Binary Independent Variable Subject to Errors of Observation," Journal of Econometrics vol. 1 (1973): 49-60.
- Angrist, Joshua D. and Alan B. Krueger. "Does Compulsory School Attendance Affect Schooling and Earnings?," Quarterly Journal of Economics vol. 106, No. 4. (1991): 979-1014.
- Ashenfelter, Orley and Alan Krueger. "Estimating the Returns to Schooling Using a New Sample of Twins," American Economic Review vol. 84, no. 4 (December 1994): 1157-1173.
- Ashenfelter, Orley and Cecilia Rouse. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins," Quarterly Journal of Economics, forthcoming 1998.
- Ashenfelter, Orley and David Zimmerman. "Estimates of the Returns to Schooling with Siblings Data: Fathers, Sons and Brothers," Review of Economics and Statistics vol. 79, No. 1 (February 1997): 1-9.
- Behrman, Jere, Zdenek Hrubec, Paul Taubman and Terence Wales, Socioeconomic Success: A Study of the Effects of Genetic Endowments, Family Environment and Schooling Amsterdam: North-Holland, 1980.
- Black, Dan A., Mark C. Berger, and Frank A. Scott, "Bounding Parameter Estimates with Non-Classical Measurement Error," University of Kentucky Working Paper, 1998.
- Bound, John and Alan B. Krueger. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," Journal of Labor Economics vol. 9, no. 1 (January 1991): 1-24.
- Butcher, Kristin and Anne Case. "The Effect of Sibling Sex Composition on Women's Education and Earnings," Quarterly Journal of Economics, vol. 109, No. 3 (August, 1994): 531-563.
- Card, David. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," Princeton Industrial Relations Section Working Paper No. 317, (July 1993).
- Card, David. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis," Econometrica, vol. 64, no. 4 (July 1996): 957-979.
- Card, David. "The Causal Effect of Education on Earnings," forthcoming in O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Vol. 3 (New York: Elsevier Science) May 1998.
- Chamberlain, Gary "Panel Data" in Z. Griliches and M. D. Intriligator (eds.), Handbook of Econometrics, Vol. 2 (New York: Elsevier Science, 1984): 1247-1318.
- Freeman, Richard B. "Longitudinal Analyses of the Effects of Trade Unions," Journal of Labor Economics vol. 2, no. 1 (January 1984): 1-26.
- Fuller, Wayne A., Measurement Error Models, (New York: Wiley, 1987).

- Griliches, Zvi. "Estimating the Returns to Schooling: Some Econometric Problems," Econometrica vol. 45, No. 1 (January, 1977): 1-22.
- Hoxby, Caroline Minter "How Teachers' Unions Affect Education Production" Quarterly Journal of Economics (1996) Vol. 111, No. 3, pp. 671-718.
- Jakubson, George. "Measurement Error in Binary Explanatory Variables in Panel Data Models: Why Do Cross Section and Panel Estimates of the Union Wage Effect Differ?," Princeton University Industrial Relations Section Working Paper No. 209 (August 1986).
- Kane, Thomas J. and Cecilia Elena Rouse. "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?," Kennedy School of Government Faculty Research Working Paper No. 93-38, (February 1993).
- Lewbel, Arthur. "Constructing Instruments for Regressions with Measurement Error When No Additional Data are Available, with an Application to Patents and R&D," Econometrica vol. 65, no. 5 (September 1997): 1201-1213.
- Mellow, Wesley, and Hal Sider. "Accuracy of Response in Labor Market Surveys: Evidence and Implications," Journal of Labor Economics vol. 1, no. 4 (January 1983): 331-344.
- Park, Jin Heum. "Returns to Schooling: A Peculiar Deviation from Linearity," Industrial Relations Section Working Paper No. 335 (October 1994).
- Rodgers III, William M. and Sarah Bruhl. "Estimating the Bias Due to Measurement Error in the Economic Returns to Schooling: Evidence from the 1990 February Current Population Survey," The College of William and Mary mimeo, August 1997.

</ref_section>

Endnotes

1. See Griliches (1977) for a survey and critique of this literature and Behrman et. al. (1980) for an example of the early research with twins.
2. See Card (1998) for a recent summary of this literature. Well known examples include Angrist and Krueger (1991); Ashenfelter and Zimmerman (1993); Butcher and Case (1992); Card (1993); and Kane and Rouse (1993).
3. Aigner (1973) considers the effect of errors of classification in binary variables on OLS estimates. See, also, Freeman (1984) who considers measurement error in a dichotomous variable in panel data.
4. For instance, Hoxby (1996) explicitly motivates her attempts to instrument for union status because the errors in reporting union status. Seemingly in confirmation of this suspicion, she finds impacts of union status much larger than OLS estimates. However, if the binary indicator of union status were measured with error, the IV estimator would be biased upward, perhaps considerably.
5. In contrast to equations (2) and (3), we include the full set of schooling categories and therefore do not include an intercept.
6. If the date of the term was missing, we assumed that it began before 1979.
7. The typical school year consists of 30 credits. Therefore, we round to the nearest 30 credits in categorizing the sample by their number of years of school completed. Specifically, we consider those with 15-45 credits as having 1 year of college, 45-75 credits 2 years of college, 75-105 credits as 3 years of college, and greater than 105 credits as four or more years of college. Those with less than 15 credits are categorized as having completed no years of college. There were two empty cells: those with self-reported 1 year of college and transcript-reported 4 years of college (no degree) and those with self-reported 4 -years of schooling and a transcript reported associate's degree. We randomly assigned 2 observations into each of these cells in order to preserve identification of the parameters in the model.
8. Of course, the extent of bias depends critically on the data being used (both the measure of education, and the instrument). For example, we experimented with an alternative specification in which we used only the credit information to assign years of education, ignoring the degree information altogether. In that case, OLS estimates of the payoff per year of schooling are biased downward, but IV estimates are biased upward by 5 to 10 percent.
9. Ninety percent of the parents were surveyed within 7 months of the student survey. Moreover, 97 percent of the parents surveyed reported that the youth had dropped out before the date of the youth's interview. As a result, any discrepancy in the student-reported and parent-reported educational attainment does not seem to be due to any difference in the timing of the interviews.
10. The questions asked of students and parents at the time of the second follow-up were quite similar: Student drop-outs were asked "When did you last attend school (a school granting or leading to a high school diploma)?" and "What grade were you in then?" Parents were asked "What is the last grade your teenager attended?"

11. The 10 students were drawn from those where the student had reported completing 12th grade or where the parent had reported that the student had completed 8th or 9th grade.

12. The results in the top panel are similar, but not identical, to the results in Angrist and Krueger (1991), Table III, Panel B. We have limited the sample to those who reported between 8 and 12 years of schooling and we have used years of schooling attended rather than years of schooling completed to be comparable to the estimates from the NELS.

Appendix Table 1
Estimates of the Measurement Error Probabilities
in Measures of Years of College and Degree Attainment

		Conditional upon Actual Schooling Level:						
		No Coll	1 Year	2 Years	3 Years	4+ Yrs	AA	BA
Transcript:	No	.948	.306	.009	.000	.312	.059	.022
	College	(.008)	(.062)	(.033)	(.011)	(.061)	(.032)	(.005)
	1	.043	.608	.090	.083	.599	.000	.002
	Year	(.007)	(.056)	(.080)	(.069)	(.054)	(.048)	(.002)
	2	.003	.074	.846	.000	.075	.171	.004
	Years	(.002)	(.036)	(.142)	(.043)	(.037)	(.044)	(.004)
	3	.002	.007	.055	.263	.008	.052	.003
	Years	(.001)	(.015)	(.058)	(.069)	(.014)	(.014)	(.002)
4+	.000	.005	.000	.616	.001	.020	.008	
Years	(.000)	(.030)	(.133)	(.137)	(.026)	(.016)	(.004)	
AA	.001	.000	.000	.038	.003	.697	.006	
Degree	(.000)	(.029)	(.057)	(.125)	(.030)	(.078)	(.003)	
BA	.001	.000	.000	.000	.001	.000	.956	
Degree	(.000)	(.006)	(.048)	(.124)	(.005)	(.016)	(.009)	
Self- Reported:	No	.985	.223	.137	.009	.225	.032	.017
	College	(.009)	(.058)	(.038)	(.054)	(.059)	(.019)	(.004)
	1	.008	.552	.033	.035	.547	.009	.002
	Year	(.006)	(.064)	(.046)	(.104)	(.064)	(.035)	(.001)
	2	.000	.164	.606	.019	.169	.017	.002
	Years	(.002)	(.038)	(.078)	(.278)	(.038)	(.035)	(.004)
	3	.001	.002	.187	.286	.002	.000	.004
	Years	(.000)	(.011)	(.038)	(.106)	(.011)	(.018)	(.004)
4+	.000	.000	.037	.652	.000	.026	.052	
Years	(.001)	(.018)	(.031)	(.240)	(.019)	(.064)	(.009)	
AA	.005	.049	.000	.000	.048	.910	.009	
Degree	(.003)	(.055)	(.082)	(.042)	(.055)	(.097)	(.004)	
BA	.000	.009	.000	.000	.009	.006	.913	
Degree	(.001)	(.008)	(.024)	(.063)	(.008)	(.014)	(.013)	
Estimated Proportion in Each Category:		.525	.038	.040	.022	.050	.062	.263
		(.012)	(.019)	(.009)	(.009)	(.018)	(.009)	(.009)

Note: Standard errors in parentheses. There were 5281 observations. The above estimates were estimated with the specification in Table 7.

Source: Authors' calculations using the NLS-72.