

IZA DP No. 1747

**School Progression and the  
Grade Distribution of Students:  
Evidence from the Current Population Survey**

Elizabeth U. Cascio

September 2005

# **School Progression and the Grade Distribution of Students: Evidence from the Current Population Survey**

**Elizabeth U. Cascio**

*University of California, Davis  
and IZA Bonn*

Discussion Paper No. 1747  
September 2005

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
Email: [iza@iza.org](mailto:iza@iza.org)

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

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

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

## ABSTRACT

### **School Progression and the Grade Distribution of Students: Evidence from the Current Population Survey\***

Education researchers have long made inferences about grade retention from the grade distribution of same-aged students. Recent economics studies have followed suit. This paper examines the validity of the “below grade” proxy for retention using data from supplemental questionnaires administered in the U.S. Current Population Survey during the 1990s. I estimate that 21% of non-repeaters are below grade, while 12% of repeaters are not. Misclassification attenuates regression coefficients by 35% when the proxy is an outcome and by 65% when it is a regressor. The latter figure is a benchmark, as classification and regression errors are arguably correlated. Biases are likely substantial in other surveys and time periods.

JEL Classification: I21, C81

Keywords: grade retention, misclassification, nonclassical measurement error

Corresponding author:

Elizabeth U. Cascio  
Department of Economics  
University of California, Davis  
One Shields Avenue  
Davis, CA 95616-8578  
USA  
Email: [eucascio@ucdavis.edu](mailto:eucascio@ucdavis.edu)

---

\* I am thankful to David Card for helpful discussions and to Ethan Lewis, Marianne Page, and Ann Huff Stevens for reading earlier drafts. All errors and omissions are my own.

## 1. Introduction

The association between grade retention and poor educational outcomes is one of the most well documented relationships in education research.<sup>1</sup> In a meta-analysis of the literature, Holmes (1989) finds that, on average, later test scores of children retained in grade are 0.19 to 0.31 standard deviations lower than those of similar children progressing normally through school. A large number of studies have also uncovered a strong association between retention and high school dropout (e.g., Grissom and Shepard, 1989; Roderick, 1994; Jimerson, 1999).

Recent quasi-experimental studies by economists suggest that these relationships may be driven by selection (e.g., Eide and Showalter, 2001; Jacob and Lefgren, 2004). Nonetheless, because grade retention is costly, policies and practices that might reduce retention alone are of great interest.<sup>2</sup> For example, many have hypothesized that early childhood programs lower the need for retention by promoting school readiness, making retention an outcome of interest in evaluations of public preschool (Cascio, 2004), Head Start (Currie and Thomas, 1995), and numerous “model” early interventions (reviewed in Barnett (1995) and Currie (2001)). Others have argued that family investments encourage normal school progression, linking retention to parental education (Oreopoulos, Page, and Stevens, 2004; Page, 2005) and sibship size (Conley and Glauber, 2005).

Despite compelling reasons to be interested its causes and consequences, grade retention is difficult to study because data are lacking in large-scale population surveys, such as the Census. The large sample sizes, consistency over time, and geographic representation of population surveys make them attractive for studies where state or local education policy changes are a source of identifying

---

<sup>1</sup> Throughout the paper, I use the terms “repetition” and “retention” interchangeably.

<sup>2</sup> Eide and Showalter (2001) estimate that the costs of grade retention range between \$2.6 and \$13 billion per year. Applying their assumptions on annual retention rates to more recent estimates of the average annual per-pupil current expenditure in the United States (\$7376 in 2000-01) and the number of students enrolled in public schools (47.2 million in 2000-01), the annual cost of grade retention could reach \$17.4 billion (in 2000 dollars).

variation (e.g., Cascio, 2004; Oreopoulos, Page, and Stevens, 2004; Page, 2005). Population surveys are also useful simply because their large sample sizes are helpful in identifying small effects.

Surveys where grade repetition can be observed or imputed—such as the cohort surveys conducted by the National Center for Education Statistics (NCES)—have relatively small samples and follow only a few cohorts, making them of limited use in these contexts. Administrative data also tend to be available sporadically and for limited subpopulations, such as individual school districts or states.

This paper examines the validity of a proxy for grade retention that can be readily constructed from population surveys—whether a child is “below grade” for his age. Grade-for-age has long been used to mark changes over time in grade progression in the United States (Rose, Medway, Cantrell, and Marus, 1983; Shepard and Smith, 1989; Roderick, 1994; Hauser, Pager, and Simmons, 2004) and has recently been used in a number of economic analyses (e.g., Cascio, 2004; Oreopoulos, Page, and Stevens, 2004; Conley and Glauber, 2005; Page, 2005). Nonetheless, no systematic evidence exists on its quality as a proxy for retention. Using data from a special battery of questions administered in the 1990s as part of the October Current Population Survey (CPS) School Enrollment Supplement, I am able to compare reported grade repetition experiences against grade-for-age in the population of school-aged students.<sup>3</sup> I then examine the consequences of misclassification. Because grade retention is a binary variable, misclassification attenuates estimates of regression parameters, whether it serves as a regressor (Aigner, 1973) or an outcome (Hausman, 2001) of interest.

I find that the extent of misclassification in the proxy and the resulting attenuation biases are considerable. Around 21 percent of non-repeaters are classified as below grade, while 12 percent of repeaters are not. Under plausible assumptions on recall bias in response to the repetition questions,

---

<sup>3</sup> Although they contain retrospective data on retention, cohort surveys conducted by the NCES, such as the High School and Beyond (HSB) and the National Education Longitudinal Study (NELS), cannot be used for a validation study, since samples are selected on children having reached a particular grade level (8<sup>th</sup> grade in the NELS and 10<sup>th</sup> grade in the HSB).

I estimate that regression coefficients will be attenuated by 35 percent when below grade is an outcome and by 65 percent when it serves as a regressor. The latter figure serves only as a benchmark, as classification and regression errors are likely to be correlated. In particular, false positives are common among children who delay school entry or reside in states with school entry cutoff dates earlier than October (the month in which age is measured), and false negatives common among children who enter school early or reside in states where entry cutoff dates are later in the academic year. Between 11 and 15 percent of misclassification is accounted for by these factors. Although the distribution of school entry dates and the propensity to delay school entry have changed over time, simulations suggest that attenuation biases are likely to remain substantial regardless of the cohorts under observation.

The remainder of the paper is organized as follows. The next section describes the October CPS data in more detail. In Section 3, I then present cross-tabulations of the two grade repetition measures and baseline estimates of the proxy's reliability. Section 4 ties misclassification rates to the factors described above and uses the resulting models to predict attenuation factors for other cohorts. Section 5 concludes.

## **2. Data and Preliminary Statistics**

### *2.1 October CPS Sample*

In 1992, 1995, and 1999, the October CPS School Enrollment Supplement included a series of non-basic questions on experiences with grade repetition. Although the universe of respondents changed from year to year, the questions applied to most young respondents currently enrolled in school. The main question asked was identical across all years in the survey.<sup>4</sup> Follow-up questions

---

<sup>4</sup> The main question asked is, "Since starting school, has [*the respondent*] ever repeated a grade?"

differed in wording from year to year, though they were substantively similar across years, asking repeaters to report the grade(s) in which they were retained.

From these files, I draw a sample of children of compulsory school age (defined here as ages 7 to 15), with non-missing data on grade repetition—roughly 96 percent of children surveyed in this age group. Since non-response rates were higher in 1992 to 1995, I have also estimated models separately for the 1999 CPS sub-sample, where response was nearly universal. In practice, the findings for this year are quite similar to those for the pooled sample and are therefore not reported in the paper.<sup>5</sup> In addition to the grade retention data, I keep information on each respondent’s enrollment status, highest grade attended, age, gender, race, and state of residence.

For the analysis, I create an indicator for retention (*repeat<sub>i</sub>*) for each individual *i* in the sample based on their answers to the main grade repetition question described above. The proxy for grade repetition (*below<sub>i</sub>*) is then constructed using data on age (*age<sub>i</sub>*), enrollment status (*enrolled<sub>i</sub>*), and grade attending (*grade<sub>i</sub>*), all measured as of the October survey:

$$(1) \quad below_i = \begin{cases} 1(age_i - grade_i > 5) & \text{if } enrolled_i = 1 \\ 1 & \text{if } enrolled_i = 0 \end{cases}$$

where  $1(\cdot)$  is the indicator function. If “on grade,” a child will start kindergarten at age five, first grade at age six, etc. Thus, equation (1) classifies a first grader who is seven years old by October as a repeater. Further, if a child is not enrolled, he is classified as below grade. Since most of the population observed will be currently enrolled in school (see Table 1), most variation in the proxy derives from variation in grade of enrollment among individuals of the same reported age.

In the next section, I compare *below<sub>i</sub>* to *repeat<sub>i</sub>* to determine the extent and consequences of misclassification. For descriptive purposes, however, I aggregate the grade repetition variables in the

---

<sup>5</sup> These results are available from the author upon request. In 1992 and 1995, data are missing for 6 and 7 percent of the school-aged population, respectively, but the non-response rate in 1999 is close to zero (see Table 1). Appendix Table 1 shows that non-response is associated with being below grade and with race.

1992, 1995, and 1999 October CPS samples to cohort/survey year level averages (where cohort is defined as survey year - age). To these averages, I then merge the cohort-level fraction below grade at different ages as observed in other years of the October CPS School Enrollment Supplement.<sup>6</sup> Although these aggregated data are not used in the formal analysis, looking across survey years provides a check on the quality of response to the grade repetition questions. This is useful, since the validation data set in this application is not from an independent (administrative) source, but rather from the same data from which the proxy itself is calculated.

## 2.2 *Summary Statistics*

Table 1 shows characteristics of the disaggregated October CPS sample by survey year and for all years combined for various population subgroups. The below grade proxy overstates the degree of grade repetition experienced in the population. In 1992, for example, slightly over 11 percent of the sample reports having repeated a grade. By contrast, nearly 28 percent are below grade, making for overstatement by more than a factor of two. The below grade measure does, however, appear to carry some signal: for example, groups with relatively high repetition rates (e.g., ages 12 to 15, males, and blacks) are also relatively more likely to be enrolled below grade.

That the proxy carries some signal is also suggested by Figure 1, which plots inter-cohort trends in the percent of children who report repeating either kindergarten or first grade (or both) in each of the 1992, 1995, and 1999 October CPS files. Alongside, I present the percent of the cohort below grade by age seven, calculated using basic October CPS education data from other years. For completeness, I extend the 1992, 1995, and 1999 samples to include reported grade repetition rates of individuals between the ages of 16 and 24 when surveyed. I also extend the below grade series to

---

<sup>6</sup> Age, grade, and enrollment status are measured as of October. All tabulations are weighted by final CPS weights. Between 1968 and 1986, approximately 0.5 percent of 7 to 15 year olds report attending a “special school” without information on a grade equivalent. There is not a strong relationship between this response and year or age. I therefore code individuals enrolled in special schools as below grade.



include cohorts for whom repetition was not reported in the 1992, 1995, and 1999 October CPS samples. Consistent with Table 1, Figure 1 shows that the below grade measure seriously overstates the proportion of children experiencing grade retention: the average difference between the two series ranges between 10 and 15 percentage points. However, the two series move together, supporting the idea that the proxy conveys information on grade repetition as well as the contention that responses to the grade repetition questions were truthful. Moreover, overall trends in both measures fit received wisdom on retention policy, suggesting that widespread social promotion ended in the early 1980s and did not resume until the 1990s (e.g., Roderick, 1994).

Subgroup-specific grade repetition rates reported in Table 1 are also largely consistent with small-scale administrative and survey data on grade retention (e.g., Rose, *et al.*, 1983, Eide and Showalter, 2001).<sup>7</sup> For example, boys are less likely to progress through school at a normal rate; about 12 percent of males between the ages of 7 and 15, compared to only 7.5 percent of females (not shown in table), report having repeated a grade. Similarly, a grade repetition experience is documented for 14.5 percent of blacks, compared to only 8.5 percent of non-Hispanic whites (not shown in table). Both males and blacks are relatively more likely to have repeated kindergarten or first grade, suggesting that these differences in grade progression arise early in the school career.

### **3. Estimates and Consequences of Misclassification**

#### *3.1 Cross-Tabulations*

Table 2 gives a cross-tabulation of  $below_i$  and  $repeat_i$  in the October CPS sample along with corresponding misclassification rates. Individuals are classified into one of four groups: (1)

---

<sup>7</sup> In terms of magnitudes, repetition rates in the October CPS are also similar to those in other surveys where repetition questions are retrospective. For example, the sample of HSB sophomores in Eide and Showalter (2001) have retention rates of 16 percent (white men), 10 percent (white women), 21 percent (black men), and 17 percent (black women). The October CPS sample of 15 year olds yields retention rates of 16.0 percent (white men), 9.6 percent (white women), 23.7 percent (black men), and 16.8 percent (black women).

repeated and currently below grade (correctly classified); (2) have not repeated and not currently below grade (correctly classified); (3) repeated and not currently below grade (not correctly classified); and (4) have not repeated and currently below grade (not correctly classified). The gross misclassification rate is therefore given by the probability of falling in one of the last two categories. The final two columns present estimates of conditional probabilities of misclassification, which are useful in the analysis below: the false negative rate, or the proportion of repeaters misclassified ( $p_0 \equiv \Pr(\text{below}_i = 0 \mid \text{repeat}_i = 1)$ ), and the false positive rate, or the proportion of non-repeaters misclassified ( $p_1 \equiv \Pr(\text{below}_i = 1 \mid \text{repeat}_i = 0)$ ).<sup>8</sup>

Consistent with the discussion above, most misclassification arises from non-repeaters. The first row of Table 2 shows that 18.6 percent of respondents are below grade, but have never been retained. This category alone accounts for 94 percent of gross misclassification. The false positive rate is, however, only about twice the size of the false negative rate: 11.8 percent of repeaters are misclassified, compared to 20.6 percent of non-repeaters. The table also shows a small amount of variation across population subgroups in the degree and direction of misclassification. For example, boys and Hispanic children are relatively more likely to be misclassified (with gross misclassification rates of 21.9 and 23.0 percent, respectively) and to be misclassified as repeaters (with false positive rates of over 23 percent). False negatives are relatively common among Asians/Pacific Islanders/ Native Americans (19.2 percent) and Hispanics (17.7 percent). In general, however, misclassification rates vary over a narrow range. Overall misclassification rates range between 17.6 percent (for girls) and 22.9 percent (for Hispanics). With the exceptions listed above, false positives range between 17.9 percent (girls) and 22.1 percent (older children), and false negatives range

---

<sup>8</sup> In Table 2, estimates of  $p_0$  and  $p_1$  are given separately by population subgroup. Strictly speaking,  $p_0$  and  $p_1$  should therefore be defined as conditional on observables. For simplicity, this additional notation is not introduced in this section.

between 10.2 percent (whites) and 13.9 percent (girls). Some across-group differences are statistically significant, as shown below.

Classification errors have implications for analyses where the below grade proxy is used as either an explanatory or dependent variable. Measurement error in a binary dependent variable will attenuate estimates of regression parameters (Hausman, 2001). Coefficient estimates also be attenuated by misclassification when the proxy is used as a regressor (Aigner, 1973). Most existing studies in economics that use the proxy have been of the former type; implications of the latter case are discussed for completeness.

### 3.2 Consequences of Misclassification: the Dependent Variable Case

Suppose that the model of interest is the linear probability model  $repeat_i = \mu + \theta' X_i + v_i$ , where  $X_i$  is a  $K \times 1$  vector of regressors, and  $X_i$  and  $v_i$  are uncorrelated.<sup>9</sup> Letting  $below_i = repeat_i + \omega_i$ <sup>10</sup> and assuming that the classification errors are independent of  $X_i$  conditional on true repetition status, then

$$(2) \quad p \lim \hat{\theta}^{OLS} = \theta(1 - p_0 - p_1),$$

where  $p_0$  and  $p_1$  are the false negative and false positive rates defined above, and  $\theta$  is the  $K \times 1$  parameter vector of interest, capturing partial relationships between each regressor and the probability of repeating a grade. OLS estimates of  $\theta$  are thus biased by a proportional factor of  $\tau \equiv 1 - p_0 - p_1$ ; any misclassification in the proxy—regardless of its direction—biases the OLS

---

<sup>9</sup> Suppose that  $y_i^* = \tilde{\theta}' X_i + \varepsilon_i$ , where  $y_i^*$  is some continuous (unobserved) index of academic performance, and that  $repeat_i = 1(y_i^* \leq 0)$ . It follows that  $E[repeat_i | X_i] \equiv \Pr(repeat_i = 1 | X_i) = F(-\tilde{\theta}' x_i)$ , where  $F(\cdot)$  is the CDF of  $\varepsilon_i$ . Researchers might assume that  $\varepsilon_i$  is uniformly distributed, so that  $F(\cdot)$  is a linear function (i.e., a linear probability model). The parameter  $\theta$  is then simply a rescaled version of  $\tilde{\theta}$ .

<sup>10</sup> Measurement error in a binary variable can undertake one of only three possible values, i.e.,  $\omega_i \in \{-1, 0, 1\}$ . Measurement error is nonclassical, as the signal,  $repeat_i$ , is necessarily (negatively) correlated with the noise,  $\omega_i$ .

estimator of  $\theta$  toward zero. A similar result holds in nonlinear models, such as probit or logit (Hausman, 2001).<sup>11</sup>

In most applications, the model of interest is one where at least one component of  $X_i$  is potentially correlated with  $v_i$  and at least one instrument is forwarded to remove the resulting bias in  $\hat{\theta}^{OLS}$ . Let  $X_i' = [\tilde{x}_i' \ x_i']$ , where  $\tilde{x}_i$  is a  $\tilde{K} \times 1$  vector of endogenous regressors ( $\tilde{K} \leq K$ ) and  $x_i$  is a  $(K - \tilde{K}) \times 1$  vector of exogenous regressors, and let  $\theta' = [\theta_{\tilde{x}}' \ \theta_x']$  represent the corresponding vector of regression parameters. With misclassification in the dependent variable,

$$(3) \quad p \lim \hat{\theta}^{OLS} = \theta \tau + \Sigma_{XX}^{-1} \Sigma_{Xv},$$

where  $\Sigma_{XX} \equiv p \lim \left( \frac{1}{n} \sum_{i=1}^n X_i X_i' \right)$  and  $\Sigma_{Xv} \equiv p \lim \left( \frac{1}{n} \sum_{i=1}^n X_i v_i \right)$ . Thus,  $\hat{\theta}^{OLS}$  may not in fact be attenuated, even with misclassification in *below*<sub>*i*</sub>. For example, if  $\theta_{\tilde{x}} > 0$  and if unobserved determinants of *repeat*<sub>*i*</sub> are positively correlated with all components of  $\tilde{x}_i$ , then  $\hat{\theta}_{\tilde{x}}^{OLS}$  will be less attenuated than if misclassification were the sole source of bias.

Instrumental variables estimators may remove endogeneity bias, but they do not remove the attenuation bias that results from misclassification in a dependent variable. Let  $Z_i$  represent a  $J \times 1$  vector of instruments ( $J \geq \tilde{K}$ ), which are correlated with the endogenous regressors,  $\tilde{x}_i$ , but not correlated with  $v_i$ . Then the two stage least squares (2SLS) estimator of  $\theta$  is attenuated by the same proportional factor as was the OLS estimator when there was no endogeneity bias, i.e.,

$$(4) \quad p \lim \hat{\theta}^{2SLS} = \theta \tau.$$

Intuitively, if  $Z_i$  predicts  $\tilde{x}_i$  and  $\tilde{x}_i$  predicts *repeat*<sub>*i*</sub> (that is, if there is a first stage and if  $\theta \neq 0$ ), then  $Z_i$  must be correlated with the classification error,  $\omega_i$ . Since  $\omega_i$  becomes subsumed in the

---

<sup>11</sup> In this case,  $E[\textit{below}_i | X_i] = p_1 + (1 - p_0 - p_1)G(\theta' X_i)$ , where  $G(\cdot)$  is the CDF of  $v_i$ .

regression error, the attenuation bias persists. The same intuition underlies the bias in OLS in the simpler case, shown in equation (2).

Thus, depending on the magnitude and sign of  $\Sigma_{XX}^{-1}\Sigma_{Xv}$ ,  $\hat{\theta}^{2SLS}$  might be more biased than  $\hat{\theta}^{OLS}$ . The October CPS sample provides a source of information on the magnitude of the bias in  $\hat{\theta}^{2SLS}$  (and in  $\hat{\theta}^{OLS}$  when there is no endogeneity bias).<sup>12</sup> Table 3 gives estimates of  $\tau$  by population subgroup; estimates of  $p_0$  and  $p_1$  are repeated from Table 2. In the pooled sample,  $\tau = 0.675$ , implying that regression coefficients will be attenuated by 32.5 percent when *below* is used as an outcome. Estimates of  $\tau$  vary across population subgroups, but over a fairly narrow range. For example, the minimum value undertaken by  $\tau$  is for Hispanic students ( $\tau = 0.588$ ), implying that regression coefficients for this subgroup will be attenuated by more than 40 percent. The maximum value undertaken by  $\tau$  is for non-Hispanic white students. Even for this subgroup, however, regression coefficients remain attenuated by over 30 percent ( $\tau = 0.698$ ).

### 3.3 *Consequences of Misclassification: the Explanatory Variable Case*

The attenuation factor assumes a different form when *below* is used as a regressor. Consider the simple bivariate model given by  $y_i = \alpha + \beta \textit{repeat}_i + u_i$ , where  $\textit{repeat}_i$  and  $u_i$  are uncorrelated. Once again letting  $\textit{below}_i = \textit{repeat}_i + \omega_i$  and assuming that  $\omega_i$  is independent of  $u_i$ ,

$$(5) \quad p \lim \hat{\beta}^{OLS} = \beta \lambda,$$

where  $\lambda \equiv \text{cov}(\textit{repeat}_i, \textit{below}_i) / \text{var}(\textit{below}_i)$  is the attenuation factor, commonly referred to as the reliability ratio (Angrist and Krueger, 1999). With a misclassified binary regressor, it is always the

---

<sup>12</sup> In nonlinear models,  $p_0$  and  $p_1$  (and therefore  $\tau$ ) can be estimated using maximum likelihood estimation techniques (Hausman, Abreyva, and Scott-Morton, 1998). However, most applications using *below* employ linear probability models (or aggregated versions thereof). If linear models are appropriate, additional information is needed to identify these parameters.

case under these assumptions that  $\lambda < 1$  (Aigner, 1973). In this application, the reliability ratio can be reduced to a rescaled version of  $\tau$ , given by

$$(6) \quad \lambda = \frac{\text{var}(\text{repeat}_i)}{\text{var}(\text{below}_i)} \tau,$$

and can be readily estimated as the slope in a regression of  $\text{repeat}_i$  on  $\text{below}_i$ . Summary statistics presented above suggest that  $\text{var}(\text{repeat}_i) < \text{var}(\text{below}_i)$ , so misclassification is likely to exert relatively more attenuation bias in applications where  $\text{below}$  is used as a regressor. Moreover, as with classical measurement error, attenuation bias is exacerbated in multivariate regression. The attenuation factor in this case is given by

$$(7) \quad \lambda_x = \left( \lambda - \frac{R^2}{1 - \rho_0 - \rho_1} \right) / (1 - R^2),$$

where  $R^2$  is from a regression of  $\text{below}$  on  $X$  (Card, 1996).

The bias term is more complicated when misclassification and regression errors are correlated (e.g., Black, Sanders, and Taylor, 2003). Such a situation is likely to arise with the proxy under consideration here. For example, as shown below, false positives are correlated with delayed school entry, which may be related to unobservables that determine high school dropout or other educational outcomes. False negatives are correlated with accelerated school entry, and the same reasoning applies. Considering the bivariate model and allowing  $\omega_i$  and  $u_i$  to be correlated,

$$(8) \quad p \lim \hat{\beta}^{OLS} = \beta \lambda + \frac{\text{cov}(\omega_i, u_i)}{\text{var}(\text{below}_i)}.$$

The sign of the asymptotic bias on  $\hat{\beta}^{OLS}$  is therefore ambiguous, even with no endogeneity bias.

In principle, this additional bias term could be responsible for the common finding that retention is negatively associated with educational attainment. In particular, suppose that educational attainment is the outcome of interest and that  $\beta > 0$ . Retention is therefore positively

associated with attainment, perhaps because it provides a second chance to learn critical math or reading concepts. If  $\text{cov}(\omega_i, u_i) = 0$ , the OLS estimator will be asymptotically biased toward zero as a result of misclassification, though still positive. However, if positive shocks to outcomes are more common among those with false negatives (e.g., early school entrants), the second bias term will be negative (i.e.,  $\text{cov}(\omega_i, u_i) < 0$ ), and it may be the case that  $p \lim \hat{\beta}^{OLS} < 0$ . Thus, even in the absence of endogeneity bias, it is possible to arrive incorrectly at the conclusion that retention leads to poor educational outcomes. In general, instrumental variables approaches will not identify  $\beta$  in this case or in the simpler case where  $\omega_i$  and  $u_i$  are uncorrelated.<sup>13</sup>

The remainder of the first panel in Table 3 gives estimates of the baseline attenuation factor,  $\lambda$ , by population subgroup. As anticipated,  $\lambda$  is much smaller than  $\tau$ . In the pooled sample,  $\lambda = 0.295$ , suggesting that the coefficient on *repeat* will be attenuated by more than 70 percent in a bivariate model where *below* is used as a proxy. Consistent with the small differences in misclassification rates documented in Table 2, estimates of  $\lambda$  also vary across population subgroups. The largest degree of attenuation is expected for Asian/Islander/Native students ( $\lambda = 0.227$ ), and the smallest degree is expected for black students ( $\lambda = 0.385$ ). These estimates should be thought of only as a benchmark; as suggested by (8), the true bias in  $\hat{\beta}_{OLS}$  will be contingent upon the application. However, the severity of the baseline attenuation bias makes it seem plausible that misclassification can generate estimates of the wrong sign.

### 3.4 *What if the Signal Isn't Observed?*

The tabulations thus far have been made under the assumption that the battery of grade repetition questions in the October CPS elicited the truth. However, respondents might

---

<sup>13</sup> See Kane, Rouse, and Staiger (1999), Black, Berger, and Scott (2000), and Black, Sanders, and Taylor (2003) for approaches to identifying or bounding regression parameters when discrete regressors are misclassified.

misinterpret the question, lie, or fail to recall whether they repeated a grade. The likelihood of small degree of “recall bias” is high, particularly for this education-related question; for example, educational attainment tends to rise with age for any given cohort, well after the cohort has finished school and well before the cohort has experienced much mortality (Card and Lemieux, 2001).

What are the implications of a small degree of recall error in response to the grade retention questions? For simplicity, suppose that if there were any error, it would yield understatement of the truth (i.e., repeaters being recorded as non-repeaters). Following Card, Hildreth, and Shore-Sheppard (2004), suppose that this understatement is captured in one parameter,  $0 < q < 1$ , defined such that  $\Pr(\text{repeat}_i = 1) = (1 - q)\Pr(\text{repeat}_i^* = 1)$ . The variable  $\text{repeat}_i^*$  represents the truth (the signal), and  $\text{repeat}_i$  denotes the response to the grade repetition question. While  $q$  cannot be identified with the available information, an estimate of the true grade repetition rate,  $\Pr(\text{repeat}_i^* = 1)$ , and estimates of  $p_0$  and  $p_1$  follow directly from any assumption on its value.<sup>14</sup>

The second half of Table 3 shows misclassification rates and attenuation factors under the assumption that  $q = 0.2$ , or that 20 percent of the population of true repeaters report not having repeated a grade. Thus, the true grade repetition rate among seven to fifteen year olds is assumed to be 11.875 percent, instead of only 9.5 percent.<sup>15</sup> The main effect of this exercise is to increase estimates of  $\lambda$ , suggesting that regression coefficients might be somewhat less attenuated than otherwise expected when repetition is an explanatory variable. The increase in  $\lambda$  is brought about

---

<sup>14</sup> In particular,  $\Pr(\text{repeat}_i^* = 1) = \Pr(\text{repeat}_i = 1)/(1 - q)$ ,  $p_0 = \Pr(\text{repeat}_i = 1 \ \& \ \text{below}_i = 0) / \Pr(\text{repeat}_i = 1)$ , and  $p_1 = (\Pr(\text{repeat}_i = 0 \ \& \ \text{below}_i = 1) - q(1 - p_0)\Pr(\text{repeat}_i^* = 1)) / \Pr(\text{repeat}_i^* = 0)$ .

<sup>15</sup> It is difficult to infer the amount of recall bias from other surveys, where questions about grade repetition are also retrospective. Administrative data also do not provide useful benchmarks, since they tend to report retention rates in particular grades, not the fraction of children of a particular age who have ever repeated a grade. This exercise is merely intended to be suggestive.



by an increase in the relative variance of the signal. The 20 percent recall bias yields slightly lower estimates of  $\tau$ , due to minor increases in the implied rate of false positives.

## 4. Explaining and Generalizing the Results

### 4.1 Sources of Misclassification

As noted above, classification error in *below* to some extent seems predictable. For example, children who begin school earlier or later than assigned by school entry regulations (“non-compliers”) should be more likely to be misclassified: non-repeaters may be below grade because of delayed school entry, and repeaters may not be classified as below grade if they entered school early. Age as of October is also a crude proxy for the year in which a child would have entered school. In some states (i.e., those with school entry cutoff dates after October), children aged five by October are permitted to enter kindergarten, while in others (i.e., those with cutoff dates prior to October), this is not the case. False negatives are likely to be common in the former, and false positives common in the latter.

Understanding whether these sources of misclassification are relevant is important for two reasons. First, if classification errors are correlated with characteristics that are generally unobservable to researchers, estimates of  $\lambda$  given above are not likely to be representative of true biases that arise in applications where *below* is used as a regressor. This possibility has already been discussed above, but no direct evidence forwarded to support it. Second, if non-compliance and the distribution of school entry laws are strong predictors of classification errors in *below*, then estimates of both  $\lambda$  and  $\tau$  may not be generalizable to other years of the CPS and other population surveys, as the distribution of school entry cutoff dates and the probability of delaying or accelerating school entry have changed over time.

Thus, consider the following regression models for the false negative and false positive rates, respectively:

$$(9) \quad p_{0iat} \equiv \Pr(\textit{below}_{iat} = 0 \mid \textit{repeat}_{iat} = 1; x_i, x_a, x_t, z_i) = \alpha_0 + \delta'_0 x_i + \kappa'_0 x_a + \pi'_0 x_t + \gamma'_0 z_i$$

$$(10) \quad p_{1iat} \equiv \Pr(\textit{below}_{iat} = 1 \mid \textit{repeat}_{iat} = 0; x_i, x_a, x_t, z_i) = \alpha_1 + \delta'_1 x_i + \kappa'_1 x_a + \pi'_1 x_t + \gamma'_1 z_i,$$

where  $i$  denotes the individual,  $a$  denotes age, and  $t$  denotes survey year. There are four vectors of characteristics included in these linear probability models: a vector of time-invariant, age-invariant individual characteristics, such as race, ethnicity, and gender ( $x_i$ ); a vector of age indicators ( $x_a$ ); a vector of survey year indicators ( $x_t$ ); and another vector of time-invariant, age-invariant characteristics, including indicators for delayed or early school entry and for the calendar month of the school entry cutoff date relevant when  $i$  would have been of age to enter school ( $z_i$ ).

Table 2 presented predictions from versions of (9) and (10) that adjusted for one subset of covariates in  $x_i$  at a time. Because the probabilities were presented separately by race, ethnicity, and gender, it was not clear whether differences in the probability of misclassification across groups were statistically significant. Table 4 presents adjusted estimates from the pooled data in a model with additional covariates. The first two columns for each dependent variable (columns (1) and (3) for the October CPS samples) include age and year fixed effects, in addition to fixed effects for race/ethnicity and gender. The table shows that Hispanic and Asian/Islander/Native American students have significantly higher likelihoods of not being below grade despite having repeated (column (1)). Hispanic and black children have significantly higher probabilities of being below grade as non-repeaters; for girls, this probability is significantly lower (column (3)). Across-group differences predicted by these adjusted models are quite similar to those implied by the unadjusted models presented in Table 2.

The next two columns for each dependent variable add the variables in  $\mathcal{X}_i$  that can be observed for respondents in the October Current Population Survey—indicators for school entry cutoff dates. Information on school entry cutoff dates was collected from archival sources and merged to individuals on the basis of cohort (defined as above as year – age in October) and state of residence.<sup>16</sup> Coefficients on the cutoff indicators should be interpreted relative to individuals in states with October cutoffs, the indicator for which has been omitted to identify the model. The estimates are consistent with expectations. For example, false positives are significantly more common for those in states with September or earlier cutoff dates and significantly less common for those in states with cutoffs after October (column (4)), and false negatives are significantly more common for children in states with late cutoff dates (column (2)). The addition of the cutoff indicators also significantly improves the explanatory power of each model. F-statistics on the joint significance of the cutoff indicators in these models are 51.28 and 146.09, respectively.

To examine the importance of delayed or early school entry to misclassification, it is necessary to use auxiliary data. Detailed information on age at school entry for a nationally representative sample is given in a module of the 1995 National Household Education Survey (NHES), a survey conducted by the NCES. Columns (5) through (7) of Table 4 present estimates analogous to those in columns (1) and (3) using a sample of seven and eight year olds from these data.<sup>17</sup> Baseline coefficients on the race/ethnicity and gender indicators in the model of false

---

<sup>16</sup> See Appendix Table 2. Dates are available for 1968 and 1970 (collected by Cascio and Lewis, 2005); 1975, 1984, 1990, 1997, and 2000 (collected by the Education Commission of the States); and 1965, 1972, and 1978 (from the *Digest of Education Statistics*). School entry cutoff dates are only updated in the years in which they are observed. Alternative interpolation strategies (e.g., applying new cutoffs at midpoints on the time intervals in which changes occur) do not change the substantive results in Table 4.

<sup>17</sup> In 1995, the NCES administered a module on early childhood education program participation to samples of preschool and primary school aged children (the Early Childhood Education Surveys). The NHES samples used in the calculation include all children aged 7 or 8 by October 31 of the year prior to the survey who attended kindergarten (98 percent of the 1991 sample and 98.4 percent of the 1995 sample). School entry compliance rates are calculated from answers to questions about compliance behavior. Tabulations are weighted by final survey weights.

negatives (column (5)) are comparable in relative magnitudes in the NHES and the October CPS. With the exception of the coefficient on gender, the same is not true in the model of false positives (column (7)). This could be a result of differences in the cohorts under observation or the relatively small sample sizes in the NHES.<sup>18</sup> Nonetheless, adding indicators for delayed or early school entry generates the anticipated predictions: false negatives are more common among those who enter school early (column (6)), while false positives are more common among those who enter school late (column (8)). Coefficients are quite large, implying that early entry raises the probability of a false negative by 31.8 percentage points, and delayed entry raises the probability of a false positive by 42.7 percentage points. While the latter estimate is statistically significant, the former is not, owing to the small number of repeaters in the NHES.

Thus, non-compliance with school entry laws and deviations of school entry cutoff dates from the month in which age is measured (October in this application) do appear to play a role in misclassification. Given that available data preclude estimation of the true models of interest (where delayed entry, early entry, and cutoff indicators are entered simultaneously), it is difficult to place a magnitude on their individual contributions. However, a lower bound on their joint effect is established by the maximum R-square from models that include observed variables in  $z_i$ , but no other covariates, estimated in each of the two surveys. An upper bound on the joint effect (assuming independence of both factors) comes from the sum of R-squares across these models in the two surveys.<sup>19</sup> By this measure, between 11 and 15 percent of the variation in false negatives and between 11 and 12 percent of the variation in false positives is explained by non-compliance

---

<sup>18</sup> Limiting the sample to seven and eight year olds (which is necessary in the NHES) does not produce these differences. Estimating the models in columns (1) to (4) on seven and eight year olds in the October CPS sample produces estimates that are similar in magnitude but less precise than those on the CPS sample of 7 to 15 year olds. (Estimates are available from the author upon request.)

<sup>19</sup> Implications of the adjusted models are similar, since non-compliance and the distribution of cutoff dates are essentially unrelated to characteristics in  $x_i$ ,  $x_a$ , and  $x_r$ .

and deviations in school entry cutoff dates from October. Non-compliance appears to carry relatively more explanatory power in both cases.

These results reinforce the conclusion that biases in models where the proxy is used as an explanatory variable may diverge in important ways from the estimates of  $\lambda$  given in Table 3. A family's decision to start a child's school career early or late is likely related to her innate ability as well as to other investments in her schooling, and information on non-compliance is not provided in the population surveys that are generally used to construct the below grade proxy. Controlling for race/ethnicity or gender is unlikely to remove this source of misclassification. As shown in the second panel of Table 4, non-compliance explains significantly more misclassification than do these easily observed individual characteristics, and little to none of the across-group differences themselves. Classification errors that arise from mismeasurement of the year in which individuals should have started school are arguably less of a problem in this regard, as there are fewer reasons to believe that they are also correlated with key unobservables. This source of misclassification can also be removed with state fixed effects, provided that school entry laws are stable over time.<sup>20</sup>

#### 4.2 *Misclassification Rates and Attenuation Biases in Other Samples*

Without further analysis, it would seem appropriate to exercise caution in generalizing the attenuation factors in Table 3 to other data sources and other cohorts. In particular, the distribution of school entry cutoff dates has changed considerably in the last 20 years, with the proportion of students exposed to early cutoff dates rising considerably from the late 1970s to the late 1990s.

---

<sup>20</sup> Existing studies have taken two different approaches to dealing with the possibility that school entry laws change over time. Using Census data, Oreopoulos, Page, and Stevens (2004) and Page (2005) construct the below grade proxy as an indicator for being below the median grade in a cell defined by state, year of birth, and quarter of birth. Cascio (2004) includes fixed effects for school entry dates in addition to state fixed effects.

Although it is impossible to confirm with the available data, delayed school entry may have become more common over time.

To get some sense of the sensitivity of attenuation factors to these sources of misclassification, Table 5 presents estimates of  $p_0$ ,  $p_1$ ,  $\tau$ , and  $\lambda$  under alternative assumptions on the distribution of the population across school entry dates (top panel) and the incidence of non-compliance (bottom panel). Underlying models of misclassification are from the even-numbered columns of Table 4. For simplicity, the average values of all other covariates are fixed at their values in the CPS and NHES samples. In addition, all calculations maintain the sample relative probabilities that non-repeaters (relative to repeaters) reside in early or late cutoff states (upper panel) or experience delayed or early school entry (lower panel).

Attenuation factors do not appear particularly sensitive to the distribution of cutoff dates around October. For example, assuming that *none* of the population resides in states with October cutoff dates, the minimum value of  $\tau$  is 0.639 and the minimum value of  $\lambda$  is 0.285. These figures are not much different than those observed within the October CPS sample (first row in upper panel; in boldface).<sup>21</sup> Similarly, assuming that *all* of the population resides in states with October cutoff, there is still substantial misclassification, with  $\tau = 0.726$  of  $\lambda = 0.310$  (last row in upper panel; in boldface). Thus, variation in school entry cutoff dates within the year appears to do little to change the substantive conclusions of this paper. This is perhaps not that surprising, since non-compliance was the relatively important of the two sources of misclassification investigated above.

Changing the incidences of delayed and early school entry has a larger impact on estimates of  $\tau$  and  $\lambda$ , but once again, the substantive conclusions of the paper are maintained. In the extreme case where all children comply with school entry guidelines (last row in lower panel; in

---

<sup>21</sup> These figures differ slightly from those in Table 3 because the estimation sample drops individuals in states that allow school districts local discretion over school entry cutoff dates.

boldface),  $\tau = 0.798$  and  $\lambda = 0.195$ . In another extreme case where 20 percent of population delays school entry and 7.5 percent of children enter early,  $\tau = 0.626$  and  $\lambda = 0.117$ . These predicted attenuation factors are not dramatically different than those in the NHES sample (first row in lower panel; in boldface), where 9.4 percent of children delayed school entry, and 2.4 percent began school early.

Of course, an exercise such as this can only be suggestive. Relationships between observables and misclassification might change over time, and regressions coefficients in the prediction equations are inconsistent if non-compliance and school entry cutoff dates are related. Moreover, implicit in these calculations is the assumption that age is recorded as of October.<sup>22</sup> On balance, however, it seems reasonable to conclude that regression coefficients, regardless of application type, will be biased in applications where the below grade measure is employed.

## 5. Conclusion

The causes and consequences of grade repetition are likely to be of growing interest among economists in the years to come. This paper has presented estimates of the reliability of the standard proxy for it, hoping to inform future analyses where grade progression is either an explanatory or dependent variable of interest.

Using data from several October CPS School Enrollment Supplements in the 1990s, I have estimated that around 20 percent of the school-aged population is misclassified by the below grade measure. False positives arise 21 percent of the time, around twice as often as false negatives.

Misclassification attenuates regression coefficients. In applications where the proxy serves as an

---

<sup>22</sup> This is essentially the case in most existing applications. For example, in papers using Census data from 1960 to 1980, age at the survey (April) and quarter of birth can be used to define age as of October (e.g., Cascio, 2004; Oreopoulos, Page, and Stevens, 2004; Page, 2005). In applications using the 1990 and 2000 Census (e.g., Conley and Glauber, 2005), age is by necessity defined in April. However, misclassification is likely to remain a serious source of bias.

outcome, regression coefficients may be attenuated by over 30 percent; instrumental variables estimates remove endogeneity bias, but not the attenuation bias associated with misclassification. Because classification errors are related to factors that are generally not observable to researchers (such as delayed or early school entry), it is more difficult to pinpoint the direction and degree of bias in applications where the proxy is an explanatory variable. In the special case where classification and regression errors are uncorrelated, the OLS coefficient on the below grade proxy may be attenuated by more than 65 percent; when classification and regression errors are correlated, it is theoretically possible that the probability limit of the OLS estimator is of the wrong sign. Although misclassification is related to the timing and degree of compliance of school entry legislation, relationships do not appear strong enough to preclude generalization of these results to applications using data from other time periods.

## References

- Aigner, Dennis J. 1973. "Regression with a Binary Dependent Variable Subject to Errors of Observation." *Journal of Econometrics* 1(1): 49-59.
- Angrist, Joshua D. and Alan B. Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application on Instrumental Variables With Moments From Two Samples," *Journal of the American Statistical Association* 87(418): 328-336.
- . 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, Volume 3A, Eds. O. Ashenfelter and D. Card. Amsterdam and New York: Elsevier Science.
- Barnett, W. Steven. 1995. "Long-term Effects of Early Childhood Programs on Cognitive and School Outcomes." *The Future of Children* 5(3): 25-50.
- Black, Dan, Mark Berger, and Frank Scott. 2000. "Bounding Parameter Estimates with Nonclassical Measurement Error." *Journal of the American Statistical Association* 95(451): 739-748.
- Black, Dan, Seth Sanders, and Lowell Taylor. 2003. "Measurement of Higher Education in the Census and Current Population Survey." *Journal of the American Statistical Association* 98(463): 545-554.
- Card, David. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64(4): 957-979.



- Card, David, Andrew K.G. Hildreth, Lara D. Shore-Sheppard. 2004. "The Measurement of Medicaid Coverage in the SIPP: Evidence from a Comparison of Matched Records." *Journal of Business and Economic Statistics* 22(4): 410-420.
- Card, David and Thomas Lemieux. 2001. "Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?" in Jonathan Gruber (ed.), *An Economic Analysis of Risky Behavior Among Youth*. (pp. 439-482). Chicago: University of Chicago Press.
- Cascio, Elizabeth. 2004. "Schooling Attainment and the Introduction of Kindergartens into Public Schools." Manuscript, University of California, Davis.
- Cascio, Elizabeth and Ethan Lewis. 2005. "Schooling and the AFQT: Evidence from School Entry Laws." IZA Discussion Paper 1481.
- Conley, Dalton and Rebecca Glauber. 2005. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order From Exogenous Variation in Fertility." NBER Working Paper 11302.
- Currie, Janet and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85(3): 341-364.
- Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15(2): 213-238.
- Education Commission of the States. 1984. *State Characteristics: Kindergartens as of July 1984*. Denver, CO: ECS.
- . 1991. *Kindergarten Entrance Ages: 1975 and 1990*. Denver, CO: ECS.
- . 1997. *State Characteristics: Kindergarten, 1997*. Denver, CO: ECS.
- . 2000. *Kindergarten: State Characteristics*. Denver, CO: ECS.
- Eide, Eric R. and Mark H. Showalter. 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." *Economics of Education Review* 20: 563-576.
- Grissom and Shepard. 1989. "Repeating and Dropping Out of School." In *Flunking Grades: Research and Policies on Retention*, eds. Lorrie A. Shepard and Mary Lee Smith, 34-63. London: Falmer Press.
- Hauser, Robert M., Devah Pager, and Solon S. Simmons. 2004. "Race-ethnicity, Social Background, and Grade Retention." In *Can Unlike Students Learn Together? Grade Retention, Tracking, and Grouping*, eds. Herbert J. Walberg, Arthur J. Reynolds, and Margaret C. Wang, 97-114. Greenwich, CT: Information Age Publishing.
- Hausman, Jerry. 2001. "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives* 15(4): 57-68.

- Hausman, Jerry A., Jason Abrevaya, and Fiona M. Scott-Morton. 1998. "Misclassification of a Dependent Variable in a Discrete-Response Setting." *Journal of Econometrics* 87(2): 239-269.
- Holmes, C.T. 1989. "Grade Level Retention Effects: A Meta-Analysis of Research Studies." In *Flunking Grades: Research and Policies on Retention*, eds. Lorrie A. Shepard and Mary Lee Smith, 16-33. London: Falmer Press.
- Jacob, Brian A. and Lars Lefgren. 2004. "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." *Review of Economics and Statistics* 86(1): 226-244.
- Jimerson, Shane. 1999. "On the Failure of Failure: Examining the Association Between Early Grade Retention and Education and Employment Outcomes During Late Adolescence." *Journal of School Psychology* 37(3): 243-72.
- National Center for Education Statistics. Various Years. *Digest of Education Statistics*. Washington, D.C.: U.S. GPO.
- Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens. 2004. "Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling," NBER Working Paper 10164.
- Page, Marianne. 2005. "Fathers' Education and Children's Human Capital: Evidence from the World War II G.I. Bill." Manuscript, University of California Davis.
- Roderick, Melissa. 1994. "Grade Retention and School Dropout: Investigating the Association." *American Educational Research Journal* 31(4): 729-59.
- Rose, Janet S., Frederic J. Medway, V.L. Cantrell, and Susan H. Marus. 1983. "A Fresh Look at the Retention-Promotion Controversy." *Journal of School Psychology* 21: 201-211.
- Shepard, Lorrie A. and Mary Lee Smith. 1989. "Introduction and Overview." In *Flunking Grades: Research and Policies on Retention*, eds. Lorrie A. Shepard and Mary Lee Smith, 1-15. London: Falmer Press.

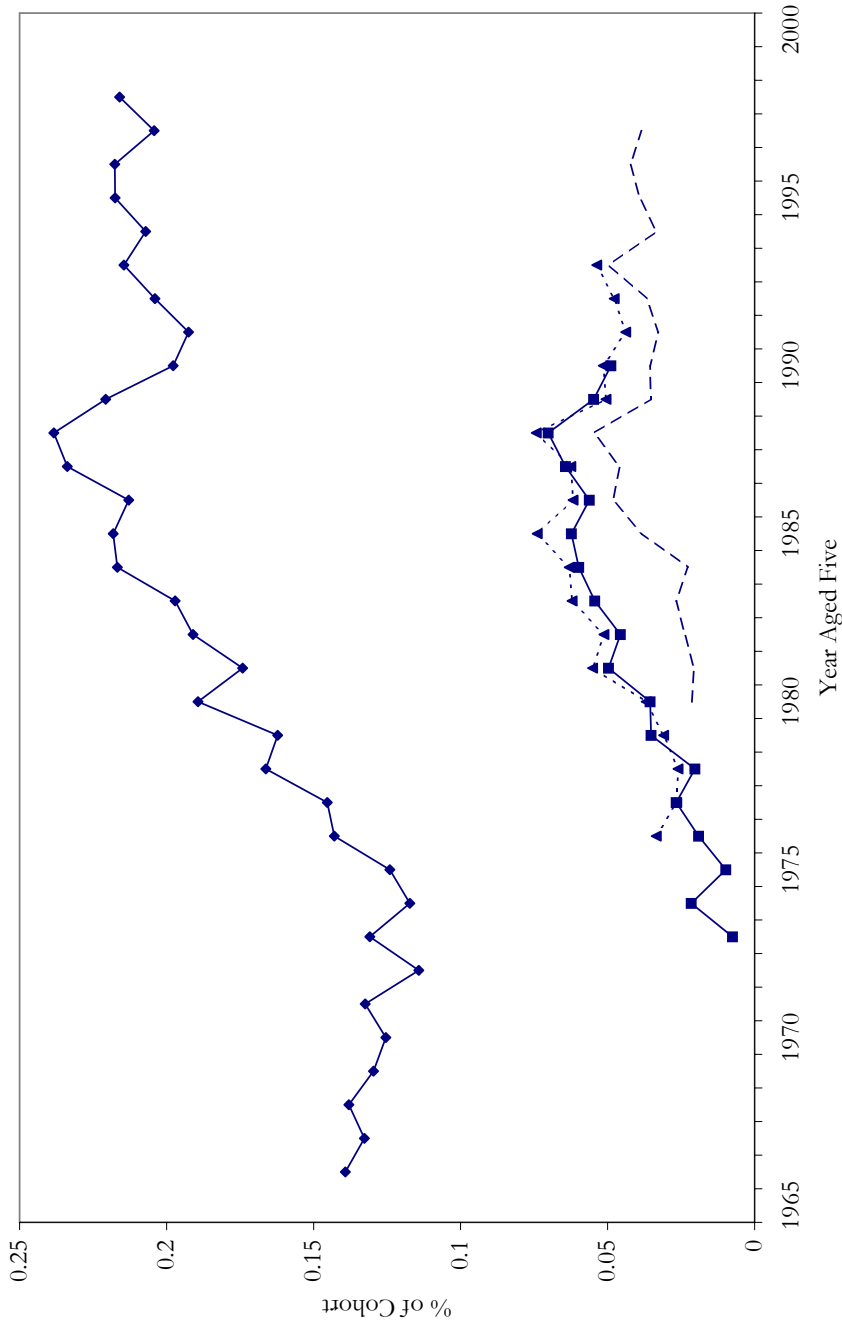


Figure 1. Inter-cohort Trends in Early Grade Progression

◆ — % below grade (at age 7)  
■ - - - repeated 1st or K, 1992  
▲ ···· repeated 1st or K, 1995

*Notes:* The percent below grade is calculated using the formula given in the text on the basis of data from the 1968 - 2000 October CPS School Enrollment Supplements. The remaining series are calculated from the battery of (retrospective) grade repetition questions administered as part of the 1992, 1995, and 1999 October CPS School Enrollment Supplements, where samples all individuals with non-missing data between the ages of 7 and 24. Year aged five is estimated as year of the survey less age at the time of the survey, plus five. All calculations are weighted by CPS final weights.

Table 1. Characteristics of School Aged Children in the October CPS: 1992, 1995, and 1999

	By Year:			All Years Combined, by Subgroup:					
	1992	1995	1999	All	Ages 12 to 15	Male	Non-Hispanic Black	Hispanic	Asian/Islander/Native American
<i>Grade Repetition Measures</i>									
Below Grade	27.7	27.8	25.9	27.1	30.3	31.1	31.3	22.9	29.3
Repeated	11.1	10.0	7.8	9.5	12.3	11.6	14.5	6.7	10.1
Repeated K	2.4	2.2	1.4	2.0	2.0	2.5	1.9	1.4	1.5
Repeated 1st	3.5	3.7	2.5	3.2	3.4	3.7	4.6	2.2	2.8
Enrolled	99.9	98.9	98.7	99.1	99.1	99.1	99.0	99.2	99.0
Grade Attending	5.7	5.7	5.7	5.7	8.2	5.7	5.7	5.8	5.6
Missing Repetition Data	6.0	6.8	0.1	4.2	4.3	4.3	5.8	5.4	4.5
<i>Other Characteristics</i>									
Age	10.9	11.0	11.0	11.0	13.5	11.0	10.9	11.0	10.9
Non-Hispanic Black	15.4	15.1	15.8	15.5	15.4	15.2	-	-	-
Asian/Islander/Native American	4.7	5.0	5.5	5.1	5.2	5.0	-	-	-
Hispanic	10.7	11.4	14.7	12.4	11.9	12.5	-	-	-
Female	49.0	48.9	49.0	49.0	49.1	-	49.7	49.3	48.9
Year=1995	-	-	-	32.7	33.2	32.7	31.9	24.1	32.7
Year=1999	-	-	-	36.4	36.3	36.4	37.4	44.9	41.4
Number of observations	18651	17161	17046	52858	23149	26776	6930	3463	5181

Notes: Author's tabulations from the 1992, 1995, and 1999 October CPS files. All figures above the line are percentages. Samples consist of children between the ages of 7 and 15 with non-missing data in the battery of grade repetition questions. All tabulations are weighted by final CPS weights.



Table 3. Estimated Attenuation Biases Resulting from Misclassification in the Below Grade Proxy

	No Recall Error ( $q=0$ )				Recall Error ( $q=0.2$ )					
	Misclassification Terms:		Bias Terms:		Misclassification Terms:		Bias Terms:			
	Not Below Grade   Repeat ( $p_0$ )	Below Grade   Not Repeat ( $p_1$ )	Relative Variance of Repeat to Below	$\tau$	$\lambda$	Not Below Grade   Repeat ( $p_0$ )	Below Grade   Not Repeat ( $p_1$ )	Relative Variance of Repeat to Below	$\tau$	$\lambda$
All years, all groups	0.118	0.206	0.44	0.675	0.295	0.118	0.211	0.53	0.670	0.357
<i>By Gender:</i>										
Female	0.139	0.179	0.39	0.682	0.264	0.139	0.182	0.47	0.679	0.322
Male	0.106	0.234	0.48	0.660	0.317	0.106	0.241	0.58	0.653	0.378
<i>By Race/Ethnicity:</i>										
Non-Hispanic White	0.102	0.200	0.40	0.698	0.283	0.102	0.204	0.49	0.694	0.343
Non-Hispanic Black	0.116	0.216	0.58	0.668	0.385	0.116	0.225	0.69	0.659	0.454
Asian/Islander/Native	0.192	0.192	0.37	0.616	0.227	0.192	0.195	0.46	0.612	0.283
Hispanic	0.177	0.236	0.44	0.588	0.256	0.177	0.242	0.53	0.581	0.309
<i>By Age Group:</i>										
Ages 7 to 11	0.127	0.196	0.37	0.677	0.250	0.127	0.199	0.45	0.673	0.304
Ages 12 to 15	0.112	0.221	0.51	0.668	0.342	0.112	0.228	0.62	0.660	0.408
<i>By Year of Survey:</i>										
1992	0.129	0.203	0.49	0.668	0.329	0.129	0.209	0.60	0.662	0.395
1995	0.116	0.211	0.45	0.673	0.301	0.116	0.216	0.54	0.667	0.363
1999	0.108	0.205	0.38	0.687	0.259	0.108	0.209	0.46	0.683	0.315

*Notes:* Author's tabulations from the 1992, 1995, and 1999 October CPS files. Samples consist of children between the ages of 7 and 15 with non-missing data in the battery of grade repetition questions. All tabulations are weighted by final CPS weights. The bias term  $\tau$  is the factor by which estimates of regression parameters will be attenuated when the below grade indicator is used as a dependent variable. The bias term  $\lambda$  is the estimate of a regression parameter will be attenuated when the below grade indicator is used as an explanatory variable, no other regressors are included in the model, and classification and regression errors are independent. For formulas, see text.

Table 4. Predictors of Misclassification

	October CPS (All Years, All Ages)				NHES (1995, Ages 7-8)			
	not below   repeat=1 (1)	repeat=1 (2)	below   repeat=0 (3)	repeat=0 (4)	not below   repeat=1 (5)	below   repeat=0 (6)	below   repeat=0 (7)	below   repeat=0 (8)
<i>Mean Dependent Variable</i>	0.121	0.121	0.201	0.201	0.078	0.078	0.192	0.192
<i>Race/Ethnicity Indicators (White Non-Hispanic Omitted):</i>								
Non-Hispanic Black	0.007 (0.014)	0.001 (0.014)	<b>0.021</b> (0.007)	<b>0.025</b> (0.007)	0.051 (0.063)	0.073 (0.070)	<b>-0.072</b> (0.023)	-0.044 (0.023)
Asian/Islander/Native	<b>0.094</b> (0.035)	<b>0.079</b> (0.034)	-0.014 (0.009)	0.005 (0.009)	0.353 (0.313)	0.268 (0.218)	<b>-0.112</b> (0.034)	<b>-0.106</b> (0.034)
Hispanic	<b>0.079</b> (0.021)	<b>0.066</b> (0.021)	<b>0.043</b> (0.008)	<b>0.057</b> (0.008)	0.116 (0.095)	0.138 (0.089)	<b>-0.072</b> (0.022)	<b>-0.069</b> (0.022)
<i>Female Indicator</i>	0.020 (0.012)	0.020 (0.012)	<b>-0.054</b> (0.005)	<b>-0.054</b> (0.005)	0.012 (0.053)	0.017 (0.051)	<b>-0.094</b> (0.016)	<b>-0.072</b> (0.016)
<i>Cutoff Indicators (October Omitted):</i>								
June-September		0.008 (0.013)		<b>0.029</b> (0.008)				
November-January		<b>0.143</b> (0.016)		<b>-0.054</b> (0.008)				
<i>Non-compliance Indicators (Complier Omitted):</i>								
Delayed Entry						0.013 (0.100)		<b>0.427</b> (0.032)
Early Entry						0.318 (0.159)		-0.001 (0.046)
R-square	0.01	0.05	0.01	0.02	0.08	0.18	0.02	0.12
F-statistic on additional covariates	-	51.28	-	146.09	-	2.01	-	86.68
Number of observations	3935	3935	38545	38545	108	108	2884	2884

*Notes:* Coefficients in bold type are significant at the 1% level. Eicker-White heteroskedasticity robust standard errors are in parentheses. Columns (1) to (4) use the subsample of respondents in states that do not allow school districts local discretion over school entry age. All regressions are weighted by survey sampling weights and include fixed effects for age. Columns (1) to (4) include survey year fixed effects. The F-statistic tests the joint significance of the cutoff indicators (columns (2) and (4)) or the non-compliance indicators (columns (6) and (8)).

Table 5. Misclassification Rates and Attenuation Biases under Alternative Assumptions on School Entry Laws and Non-compliance

		Misclassification Rates		Attenuation Factors	
		$\hat{p}_0$	$\hat{p}_1$	$\tau$	$\lambda$
<i>Cutoff Date (Frac. of Pop.):</i>					
<i>Before October</i>	<i>After October</i>	<i>Predictions based on October CPS (All years, All ages)</i>			
<b>0.485</b>	<b>0.415</b>	<b>0.122</b>	<b>0.201</b>	<b>0.678</b>	<b>0.299</b>
0	1	0.193	0.149	0.658	0.339
0.1	0.9	0.181	0.180	0.639	0.300
0.2	0.8	0.169	0.166	0.665	0.324
0.3	0.7	0.157	0.182	0.661	0.308
0.4	0.6	0.145	0.183	0.672	0.311
0.5	0.5	0.133	0.192	0.675	0.304
0.6	0.4	0.121	0.199	0.680	0.301
0.7	0.3	0.109	0.208	0.683	0.296
0.8	0.2	0.097	0.216	0.687	0.292
0.9	0.1	0.084	0.224	0.692	0.289
1	0	0.072	0.233	0.695	0.285
<b>0</b>	<b>0</b>	<b>0.064</b>	<b>0.210</b>	<b>0.726</b>	<b>0.310</b>
<i>Non-compliance (Frac. of Pop.):</i>					
<i>Delayed</i>	<i>Early</i>	<i>Predictions based on NHES (1995, Ages 7-8)</i>			
<b>0.094</b>	<b>0.024</b>	<b>0.078</b>	<b>0.192</b>	<b>0.730</b>	<b>0.155</b>
0.20	0.05	0.108	0.238	0.654	0.122
0.20	0.075	0.136	0.238	0.626	0.117
0.15	0.0375	0.094	0.216	0.690	0.136
0.15	0.05625	0.115	0.216	0.669	0.132
0.12	0.03	0.085	0.203	0.711	0.146
0.12	0.045	0.102	0.203	0.695	0.142
0.10	0.025	0.080	0.194	0.726	0.152
0.10	0.0375	0.094	0.194	0.712	0.150
0.08	0.02	0.074	0.186	0.740	0.160
0.08	0.03	0.085	0.186	0.729	0.158
0.06	0.015	0.068	0.177	0.755	0.168
0.06	0.0225	0.077	0.177	0.746	0.166
<b>0</b>	<b>0</b>	<b>0.051</b>	<b>0.151</b>	<b>0.798</b>	<b>0.195</b>

Notes: Underlying regression models are in columns (2) and (4) of Table 4 (upper panel) and in columns (6) and (8) of Table 4 (lower panel). All predictions hold constant at sample values the fractions of individuals in each race/ethnicity, gender, age, and survey year category. Also held at sample values are the *relative* probabilities that non-repeaters (relative to repeaters) reside in early or late cutoff states (upper panel) or experience delayed or early school entry (lower panel). The first line in bold type in each panel gives predictions using sample averages in each cutoff category (upper panel) or in each non-compliance category (lower panel).



Appendix Table 1. Selection into Non-Response to the Battery of Grade Repetition Questions

	Coefficient on Missing Repetition Indicator (Standard Error)									
	By Year:									
	1992	1995	1999	All	Ages 12 to 15	Male	Non-Hispanic Black	Asian/Islander/Native American	Hispanic	
Below Grade	<b>0.0409</b> (0.0049)	-0.0035 (0.0049)	<b>0.0047</b> (0.0012)	<b>0.0138</b> (0.0023)	<b>0.0153</b> (0.0034)	<b>0.0074</b> (0.0030)	<b>0.0203</b> (0.0071)	0.0045 (0.0067)	0.0181 (0.0110)	
Non-Hispanic Black	0.0102 (0.0057)	<b>0.0506</b> (0.0076)	0.0009 (0.0012)	<b>0.0207</b> (0.0032)	<b>0.0171</b> (0.0048)	<b>0.0227</b> (0.0045)	-	-	-	
Asian/Islander/Native	<b>0.0262</b> (0.0096)	<b>0.0557</b> (0.0097)	<b>-0.0009</b> (0.0003)	<b>0.0264</b> (0.0043)	<b>0.0194</b> (0.0066)	<b>0.0253</b> (0.0063)	-	-	-	
Hispanic	0.0192 (0.0081)	0.0068 (0.0070)	0.0000 (0.0009)	<b>0.0080</b> (0.0032)	0.0019 (0.0050)	0.0088 (0.0046)	-	-	-	
Female	-0.0023 (0.0039)	-0.0009 (0.0043)	-0.0001 (0.0006)	-0.0011 (0.0019)	-0.0035 (0.0029)	-	-0.0037 (0.0061)	-0.0110 (0.0060)	-0.0013 (0.0086)	
Fraction Missing Repetition	0.060	0.068	0.001	0.042	0.043	0.043	0.058	0.054	0.045	
N	19810	18358	17064	52858	24209	28019	7366	5408	3716	

*Notes:* Coefficients in bold type are significant at the 1% level. Eicker-White heteroskedasticity robust standard errors are in parentheses. All regressions are weighted by final CPS weights. Underlying data are drawn from the 1992, 1995, and 1999 October CPS files. Samples consist of children between the ages of 7 and 15.

Appendix Table 2. Cutoff Dates for School Entry, by Year and Source

State	Year and Source				
	1955 A&K (1992)	1965 DES (1967)	1968 C&L (2005)	1970 C&L (2005)	1972 DES (1973)
AL	10/1	10/1	10/1	10/1	10/1
AK	-	11/2	-	-	11/2
AZ	10/1	1/1	1/1	1/1	1/1
AR	10/1	10/1	10/1	10/1	10/1
CA	12/1	12/1	12/1	12/1	12/1
CO	9/1	-	9/1	9/1	-
CT	1/1	-	1/1	1/1	-
DE	9/1	-	9/1	1/1	-
DC	-	11/1	-	-	11/1
FL	1/1	-	1/1	1/1	-
GA	LEA	-	LEA	LEA	-
HI	-	12/31	-	-	12/31
ID	10/16	10/16	10/16	10/16	10/16
IL	12/1	12/1	12/1	12/1	12/1
IN	LEA	-	LEA	LEA	-
IA	9/15	9/15	9/15	9/15	9/15
KS	9/1	9/1	9/1	9/1	9/1
KY	10/1	12/31	12/31	12/31	12/31
LA	12/1	12/31	12/31	12/31	12/31
ME	10/15	10/15	10/15	10/15	10/15
MD	9/1	-	9/1	9/1	-
MA	LEA	-	LEA	LEA	-
MI	9/1	9/1	12/1	12/1	9/1
MN	LEA	-	9/1	9/1	-
MS	1/1	1/1	1/1	1/1	1/1
MO	10/1	10/1	9/30	9/30	10/1
MT	9/10	-	9/1	9/1	-
NE	10/15	10/15	10/15	10/15	10/15
NV	12/31	12/31	12/31	12/31	12/31
NH	9/13	-	9/13	9/13	-
NJ	10/1	-	10/1	10/1	-
NM	1/1	-	1/1	1/1	-
NY	12/1	12/1	12/1	12/1	12/1
NC	10/1	10/1	10/1	10/15	10/1
ND	10/31	10/31	10/31	10/31	10/31
OH	9/13	-	10/31	10/31	-
OK	11/1	11/1	11/1	11/1	11/1
OR	9/1	11/15	11/15	11/15	11/15
PA	2/1	2/1	LEA	LEA	2/1
RI	12/31	-	12/31	12/31	-
SC	9/1	-	11/1	11/1	-
SD	9/1	11/1	10/31	10/31	11/1
TN	12/31	12/31	9/30	9/30	12/31
TX	9/1	9/1	9/1	9/1	9/1
UT	9/2	-	9/2	9/2	-
VT	9/1	1/1	9/1	9/1	1/1
VA	9/30	9/30	9/30	9/30	9/30
WA	LEA	-	LEA	LEA	-
WV	11/1	11/1	11/1	11/1	11/1
WI	12/31	12/1	12/31	12/31	12/1
WY	9/15	9/15	9/15	9/15	9/15

Notes: A&K stands for Angrist and Krueger, C&L stands for Cascio and Lewis, DES stands for Digest of Education Statistics, and ECS stands for Education Commission of the States. See References.

Appendix Table 2. Cutoff Dates for School Entry, by Year and Source (Continued)

State	Year and Source					
	1975 ECS (1991)	1978 DES (1984)	1984 ECS (1984)	1990 ECS (1991)	1997 ECS (1997)	2000 ECS (2000)
AL	10/1	-	10/1	10/1	9/1	9/1
AK	11/2	11/2	11/2	11/2	8/15	8/15
AZ	1/1	1/1	9/1	9/1	9/1	9/1
AR	10/1	10/1	10/1	10/1	8/1	9/15
CA	12/1	-	12/1	12/1	12/2	12/2
CO	-	-	LEA	LEA	LEA	LEA
CT	-	1/1	1/1	1/1	1/1	1/1
DE	1/31	-	1/1	1/1	8/31	8/31
DC	-	12/31	12/31	12/31	12/31	12/31
FL	1/1	1/1	9/1	9/1	9/1	9/1
GA	9/1	12/1	9/1	9/1	9/1	9/1
HI	12/31	12/1	12/31	-	12/31	12/31
ID	-	10/16	10/15	8/16	9/1	9/1
IL	12/1	12/1	12/1	9/1	9/1	9/1
IN	LEA	-	LEA	LEA	6/1	6/1
IA	9/15	-	9/15	9/15	9/15	9/15
KS	9/1	9/1	9/1	9/1	8/31	8/31
KY	12/31	-	10/1	10/1	10/1	10/1
LA	1/1	1/1	1/1	1/1	9/30	9/30
ME	10/15	10/15	10/15	10/15	10/15	10/15
MD	-	12/31	12/31	12/31	12/31	12/31
MA	LEA	9/1	LEA	LEA	LEA	LEA
MI	12/1	12/1	12/1	12/1	12/1	12/1
MN	9/1	9/1	9/1	9/1	9/1	9/1
MS	-	-	-	9/1	9/1	9/1
MO	9/30	10/1	9/30	7/1	7/1	7/1
MT	9/1	-	9/10	9/10	9/10	9/10
NE	10/15	10/15	10/15	10/15	10/15	10/15
NV	-	9/30	9/30	9/30	9/30	9/30
NH	LEA	9/30	LEA	LEA	LEA	LEA
NJ	-	-	LEA	LEA	LEA	LEA
NM	11/1	9/1	9/1	9/1	9/1	9/1
NY	12/1	-	12/1	12/1	12/1	12/1
NC	10/16	10/1	10/16	10/16	10/16	10/16
ND	9/30	10/1	8/31	8/31	8/31	8/31
OH	9/30	9/13	9/30	9/30	9/30	9/30
OK	11/1	11/1	9/1	9/1	9/1	9/1
OR	11/15	11/15	11/15	9/1	9/1	9/1
PA	LEA	9/1	LEA	LEA	LEA	LEA
RI	12/31	12/31	12/31	9/30	12/31	12/31
SC	11/1	11/1	11/1	11/1	9/1	9/1
SD	10/31	11/1	9/1	9/1	9/1	9/1
TN	9/1	10/31	10/31	9/30	9/30	9/30
TX	9/1	9/1	9/1	9/1	9/1	9/1
UT	-	-	9/30	9/30	9/2	9/2
VT	12/1	-	LEA	LEA	1/1	1/1
VA	1/1	12/31	9/30	9/30	9/30	9/30
WA	9/1	11/1	8/31	8/31	8/31	8/31
WV	9/1	11/1	8/31	9/1	9/1	9/1
WI	12/1	12/1	9/1	9/1	9/1	9/1
WY	9/15	9/15	9/15	9/15	9/15	9/15