NBER WORKING PAPER SERIES

SCHOOL ENTRY, EDUCATIONAL ATTAINMENT AND QUARTER OF BIRTH: A CAUTIONARY TALE OF LATE

Rashmi Barua
Kevin Lang

Working Paper 15236
http://www.nber.org/papers/w15236

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2009

We are grateful to Josh Angrist, Sandy Black, Jim Heckman, Caroline Hoxby, Claudia Olivetti, Daniele Paserman and participants in seminars at Boston University, University of Chicago, MIT, UC Irvine, NYU, Pomona College, Tilburg University, the Tinbergen Institute, SOLE and Econometric Society for helpful comments and suggestions. The usual caveat applies. Barua acknowledges funding under NSF-AERA grant REC-0634035. Lang acknowledges funding from NICHD under grant number R03 HD056056-01 and NSF under grant number SEC-0339149. The opinions expressed here are those of the authors and not necessarily those of the funding agencies or of the National Bureau of Economic Research.

School Entry, Educational Attainment and Quarter of Birth: A Cautionary Tale of LATE
Rashmi Barua and Kevin Lang
NBER Working Paper No. 15236
August 2009
JEL No. C21,I20,J24

ABSTRACT

Partly in response to increased testing and accountability, states and districts have been raising the minimum school entry age, but existing studies show mixed results regarding the effects of entry age. These studies may be severely biased because they violate the monotonicity assumption needed for LATE. We propose an instrument not subject to this bias and show no effect on the educational attainment of children born in the fourth quarter of moving from a December 31 to an earlier cutoff. We then estimate a structural model of optimal entry age that reconciles the different IV estimates including ours. We find that one standard instrument is badly biased but that the other diverges from ours because it estimates a different LATE. We also find that an early entry age cutoff that is applied loosely (as in the 1950s) is beneficial but one that is strictly enforced is not.

Rashmi Barua
School of Economics
90 Stamford Road
Singapore 178903
rashmibarua@smu.edu.sg

Kevin Lang
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
lang@bu.edu

1 Introduction

Over the last four decades many states and school districts have increased the minimum age at which children may enter kindergarten. In the 1960s children frequently entered kindergarten when they were considerably less than five years old (or first grade when they were less than six years old). This was formally permitted in many states. In others, it was relatively easy to get around the rules. Today, thirty-eight states have cut-off dates requiring children entering kindergarten to be five years old before October 16 of the year in which they enter kindergarten, and some of the remaining states have districts that apply a stricter standard and the process continues. In recent years, Rhode Island, Maryland, Ohio and Hawaii revised their statutes to raise the kindergarten entry age. Kindergarten teachers press for stricter standards, and schools and districts, under increasing pressure to adhere to strict standards, are receptive to their arguments. Although the No Child Left Behind Act (NCLB) does not directly address the issue of entry age, one strategy for improving performance on exams is to delay the age at which children take them. Children who enter school when they are older generally do better on exams in the early elementary years. Thus, delaying the start of schooling can help states and districts meet the requirements of NCLB. However, this trend might be harmful rather than beneficial. Some critics of NCLB would argue that delayed school entry is one of its adverse unintended consequences. They argue that delayed entry reduces educational attainment and contributes to the dropout rate, particularly among children from disadvantaged socio-economic backgrounds.[1]

Whether delayed entry improves or worsens education outcomes is controversial. Following Angrist and Krueger's (1991, 1992) seminal papers, much research in both education and economics has been devoted to obtaining consistent estimates of the effects of school entry age. Angrist and Krueger address the potential endogeneity of entry age by using quarter of birth as an instrument for entry age. They show that historically individuals born in the first quarter started school later than those born in the fourth quarter, completed less education and earned less than those born in the rest of the year.[2]

Critics of this approach argue that quarter of birth may be directly related to student outcomes or parental socioeconomic status.[3] In a recent paper, Buckles and Hungerman (2008) provide evidence that children born at different times in the year are conceived by women with different socioeconomic characteristics. To address this problem, several researchers have exploited the variation in state laws governing entry age to identify its

[1] Deming and Dynarski (2008) discusses the "graying" of kindergarten and reviews elements of this debate.
[2] See also Mayer and Knutson, 1999 and Cahan and Cohen,1989.
[3] Bound, Jaeger and Baker, 1995; Bound and Jaeger, 2000.

effect on test scores, wages, educational attainment and other outcomes.[4] However, since entry age depends on both state law and date of birth, the potential endogeneity of date of birth remains problematic for this approach.

The aim of this paper is threefold. First, we address certain under-appreciated issues in the instrumental variable literature. Imbens and Angrist (1994), Angrist and Imbens (1995) and Angrist, Imbens and Rubin (1996) show that with heterogeneous treatment effects, under certain conditions, IV identifies the Local Average Treatment Effect (LATE). One condition, termed "monotonicity," generally treated as an unimportant regularity condition, requires that while the instrument may have no effect on some individuals, all of those who are affected should be affected unidirectionally. We argue that both standard instruments, quarter of birth and legal entry age, provide inconsistent estimates of LATE because they violate monotonicity. Therefore, we propose an instrument that satisfies monotonicity and gives consistent estimates of the LATE of school entry age on educational attainment.

Second, we are interested in the broader question of the optimal age at which to start school and, in particular, optimal policy regarding school entry age. It is not clear that either of the two standard instruments, even if consistent, would estimate a LATE that is of policy interest. If the law were uniform and strictly enforced and therefore monotonicity satisfied, the "born in first quarter" instrument could only hope to identify the effect of entering school when roughly six months older (on average) than those born in the other three quarters.[5] A practical policy might allow children born in the first quarter to enter a year earlier than policy previously permitted. Unless we believe that the effect of entry age is linear, the effect of an average six-month difference in entry may be very uninformative about the effect of entering a full year earlier. Assessing the LATE measured when we use legal entry age as an instrument is more complex but similar. Our instrument measures the effect on children who would otherwise enter kindergarten in the year they turn five of delaying entry until the year in which they turn six.

Third, we use simulated data to study the effect of having strict academic standards, as is practiced in schools today, on average educational attainment. Legislations, such as the No Child Left Behind (NCLB) act of 2001, have put great pressure on schools to adhere to strict academic standards. To achieve the objectives of the legislation within the specified time frame, schools are introducing academic reforms as early as preschool. Increasing academic curriculum in kindergarten is often cited as one of the reasons for rising school entry age (Deming and Dynarski, 2008; Stipek, 2006). In the past, concerns have been raised about

[4] Allen and Barnsley, 1993; Bedard and Dhuey, 2006, 2008; Cascio and Lewis, 2006; Datar, 2005; Dobkin and Ferreira, 2007; Elder and Lubotsky, 2009; Fertig and Kluve, 2005; Fredriksson and Ockert, 2006; McCrary and Royer, 2006; Puhani and Weber, 2007.

[5] Literally, the weighted average of the effect of different entry age discrepancies with a mean discrepancy of about six months.

the effect of increasing academic standards, however, to our knowledge, no study has directly tested this. Given the historical nature of our data, the instrumental variables approach captures the effect of delaying school entry as it was practiced in the 1950s. But school entry age laws are now enforced much more strictly. Consequentially, we construct a model of optimal school entry age that is well-suited to conducting several policy experiments including examining the effect of stricter enforcement. The policy experiment suggests that, in an environment where laws are strictly enforced, constraining fourth quarter children to enter late hurts these children and reduces average educational attainment.

Our two-sample two-stage least squares (TS2SLS) results show that using the quarter of birth instrument yields severely biased estimates of the effect of requiring students to enter school later than they would otherwise have chosen. When we use the consistent estimator that meets the monotonicity requirement and measures this LATE, the effect of school entry age on educational attainment is very close to zero.

However, comparing the different IV estimates does not tell us whether they diverge because the traditional estimators are inconsistent or because they are measuring different LATEs. Therefore, we develop a model of optimal school entry ages. We estimate the parameters of the model using indirect inference.

Our results show that in states with a late cutoff, the earliest optimal entry age ranges from age 4 to 4.5 years depending on the model.[6] In states with a fourth quarter cutoff the optimal entry age distribution is shifted by less than 0.01 years, and we cannot reject the hypothesis that optimal entry age is unaffected by the cutoff date.

Further, we use our simulated data to obtain "true" and IV estimates of the local average treatment effect of school entry age on educational attainment. We find that the IV estimate based on the "legal entry age" instrument is a badly biased estimator of the LATE it is intended to estimate. In contrast, in our simulations, using "born in first quarter" as an instrument generates estimates that diverge only modestly from the LATE the estimator seeks to determine.

The next section explores the literature on school entry age. Section III outlines the TS2SLS methods that we use for our baseline model. Section IV describes the data. We present the TS2SLS results in section V. Section VI builds and estimates a model of optimal school entry age. We use this model to evaluate two standard IV estimators found in the literature. In addition, we conduct policy experiments to understand the effect of different policy regimes on educational attainment. Section VII concludes.

[6] We impose that the latest optimal entry age is seven.

2 School Entry Age: Background

2.1 Literature

There has been a recent explosion of interest in school entry age that makes it difficult to treat the literature with justice. Until the 1990s, studies that looked at the effect of school entry age on student outcomes largely ignored the potential endogeneity of entry age. However, affluent parents can afford child care costs associated with delaying their child's school entry and are therefore more likely to do so. Thus, there is a positive association between parental socioeconomic conditions and entry age that can bias the OLS estimate towards a positive effect of entry age on academic outcomes. On the other hand, the OLS estimate could be downward biased if children who are less precocious intellectually and/or emotionally are redshirted[7] since these children are more likely to perform poorly on cognitive tests

Angrist and Krueger (1991, 1992), Cahan and Cohen (1989) and Mayer and Knutson (1999) address endogeneity by using quarter or month of birth as an instrument for entry age. More recent papers (Bedard and Dhuey, 2006; Datar, 2005; Elder and Lubotsky, 2009; McCrary and Royer, 2006) have used legal entry age as an instrument. This approach instruments actual entry age with the age at which the child could first legally enter school. It thus relies on both variation in state (or country) laws and month of birth.

Although somewhat mixed, the evidence from this literature suggests that older entrants have higher test scores compared to early entrants in the *same grade* and are less likely to repeat grades. However, the test score differences fade by the time the child is in middle school. Black, Devereux and Salvanes (2008) find a small beneficial effect of early entry on cognitive score at age 18. Comparing younger children of the *same age*, Barua and Lang (2008) find that early entrants perform better on achievement tests, presumably because they have completed more schooling relative to those who began school late.

Therefore it is important to determine whether entry age affects ultimate educational attainment. If late entry reduces grade retention, has no negative effect on performance within grade and has no adverse effect on ultimate grade completion, then later entry produces the same outcome at lower cost to the public (although parents pay more for child-care and their children enter the labor market later). However, if later entry is not offset by later exit, those who enter late leave school with less education (Angrist & Krueger, 1991) and fewer skills than earlier entrants leaving at that age. In this case, delaying entry reduces human capital accumulation.

The literature on the effect of entry age on educational attainment provides mixed

[7] In this context, redshirting refers to the practice of postponing entrance into kindergarten of age-eligible children in order to allow extra time for socioemotional, intellectual, or physical growth.

results. For the U.S., Angrist and Krueger (1992) and Dobkin and Ferreira (2007) find that older entrants attain slightly less education and Deming and Dynarski (2008) attribute much of the decline in educational attainment to the trend towards later school entry, but Bedard and Dhuey (2008) find no effect. Outside the US, some studies find a negative impact of early school entry on adult educational attainment and other outcomes (Allen and Barnsley, 1993; Fredriksson and Ockert, 2006) while others find positive or no effects (Fertig and Kluve, 2005; Black, Devereux and Salvanes, 2008).[8] In this paper, we argue that these findings are suspect because of important issues with the identification strategies used in the existing literature.

2.2 Specification Issues

Historically, economists assumed that instrumental variables estimates captured a single coefficient, the common effect of the explanatory variable on the dependent variable. Lang (1993) criticized the use of quarter of birth as an instrument for education in Angrist and Krueger (1991). He argued that the relation between log wage and education was inherently nonlinear and that the standard log wage equation should be viewed as a linear approximation in which the coefficient on schooling is random. He further argued that the Angrist/Krueger estimate of the return to schooling could be a severely biased estimate of the average of this random coefficient (now termed the Average Treatment Effect) because the justification for their instrument implies that it estimates the return to education only for those with relatively little education.

Building on Yitzhaki's (1989) insight that an OLS slope can be interpreted as a weighted average of adjacent slopes, Imbens and Angrist (1994) show that in a random coefficients model (i.e. one with heterogeneous treatment effects), under certain conditions, instrumental variables can still be interpreted as a Local Average Treatment Effect. LATE is the average effect of a treatment on those individuals whose treatment status is changed by the instrument. With a binary treatment, these are the "compliers" who receive the treatment when the instrument applies but not otherwise. With a multi-valued treatment, these are the individuals who increase the intensity of their treatment, and the estimated LATE gives more weight to individuals with larger responses to the instrument (Angrist and Imbens, 1995).

One of the assumptions for the identification of LATE is monotonicity: while the instrument may have no effect on some individuals, all of those who are affected must be affected in the same direction. Both the quarter of birth instrument and the legal entry age

[8] Consistent with finding no effect on educational attainment, Black, Devereux and Salvanes find a positive effect of early entry on earnings for younger workers, presumably because they have more experience. However Bedard and Dhuey find an adverse effect on earnings in the United States.

instrument violate the monotonicity assumption.[9] Many parents do not enroll their children at the earliest permissible entry age (and some find ways to enroll them earlier than is formally allowed). Such strategic behavior is more common among parents of children born in the latter half of the year (West, Meek and Hurst, 2000). Thus almost all students born in May enter kindergarten in September following their fifth birthday (or first grade following their sixth birthday). In contrast, some children born in October will enter before their fifth birthday, when they are younger than those born in May, while others will enter the following year when they are older than entrants born in May. Therefore quarter of birth is not monotonically related to school entry age.

Table 1 provides evidence of redshirting in the U.S. in the early 1950s. Using reported age and grade at the time of the 1960 Census, it shows the distribution of entry age by quarter of birth for the 1952 and 1953 cohorts. For example, among those born in 1952, 12.83% of children born in the first quarter entered kindergarten in the year that they turned four, that is when they are approximately 4.5 years old. It is evident that the proportion of children who enter kindergarten on or later than the year they turn six is highest among the fourth-quarter births. Almost 46% of 1952 born children and 40% of the 1953 cohort entered kindergarten later than age 5. So relative to other children, those born in the fourth quarter have a higher probability of entering school when they are old. However the youngest children entering school, those who are only 3.75 years old, are also born in the fourth quarter.[10]

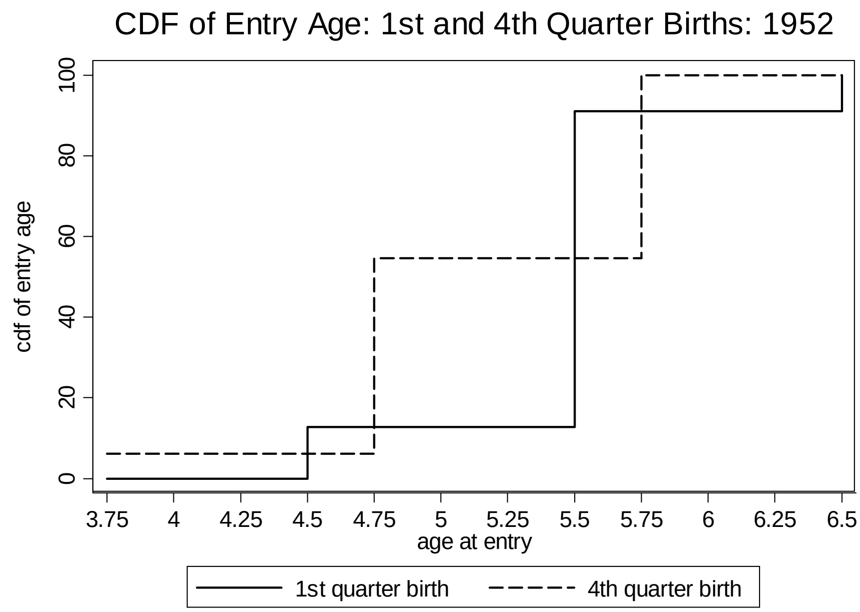
Table 1: Entrance Age by QOB and Cut Off Date

	Quarter of Birth					
	1	**2**	**3**	**4**	**4 Unconstrained State**	**4 Constrained State**
Birth Year: 1952						
4 yrs or younger	12.83	8.37	8.29	6.14	9.8	5.3
5 yrs	78.16	80.71	79.3	48.4	67.35	44.01
6 yrs and older	9.01	10.92	12.41	45.46	22.86	50.69
Birth Year: 1953						
4 yrs or younger	13.41	8.2	8.91	6.57	10.04	5.48
5 yrs	83.79	85.7	85.7	53.88	81.21	45.27
6 yrs and older	0	2.93	2.56	18.26	2.16	23.33
Not yet entered	2.8	3.17	2.83	21.29	6.59	25.92

[9] Heckman, Urzua and Vytlacil (2006) note that "uniformity" is a better term for the monotonicity condition introduced by Imbens and Angrist (1994) as the condition requires that the instrument affects all individuals in the same way rather than any particular individual in a specific way.

[10] The ages refer to the year before first grade for those who do not attend kindergarten. Equivalently, we assume that students who enter school in first grade would have spent one year in kindergarten had they enrolled. The dating of kindergarten entry is imperfect because we do not have data on retention. The higher rate of "late" entry in 1952 probably reflects the greater time that the older children have had to be retained in grade.

We summarize the comparison between children born in the first and fourth quarters in 1952 in figure 1. We can see that neither distribution of entry age is greater than the other in the sense of first-order stochastic dominance. Being born in the first quarter rather than the fourth quarter raises entry age for some children and lowers it for others.



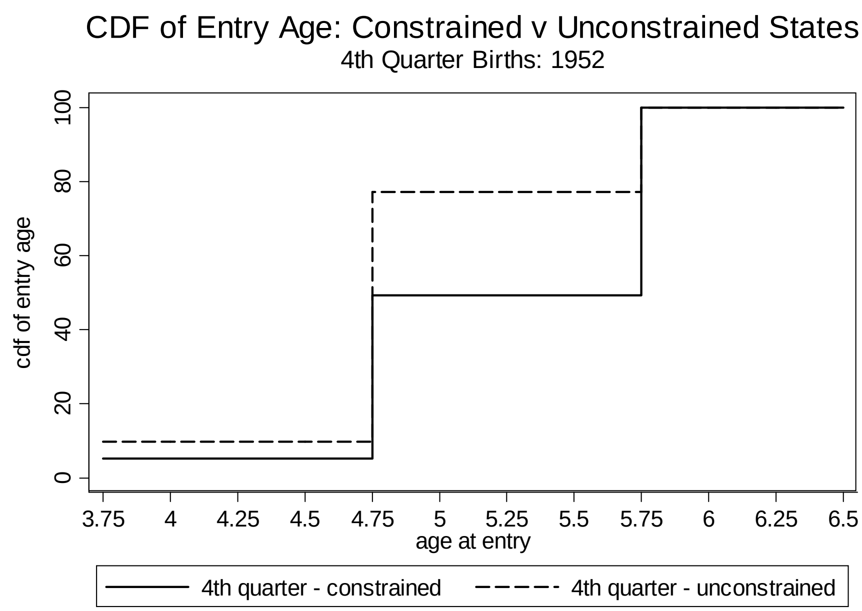
The fifth column of table 1 reports the entry age distribution for fourth quarter born children who were born before the state cutoff and were therefore eligible to enter kindergarten even before they turn five.[11] In this group, 23% of the 1952 cohort and 9% of the 1953 cohort had delayed entry. Even taking into account that some children may have been retained in grade, it is evident that there was considerable redshirting.

It is also evident that in the 1950s, there was considerable flexibility around state cutoffs. About half of children born in the fourth quarter in states with cutoffs on or before October 1 nevertheless entered school before they turned five.

It is easy to develop examples where the IV estimate using QOB or legal entry age gives severely biased estimates because of the failure of the monotonicity assumption. For simplicity assume that children are born either in the first half of the year, in which case they enter school when they are 5.5 years old or in the second half of the year in which case they can enter when they are either 5 or 6. Suppose that all children benefit from entering school when they are older. However, 75% get a benefit of 4 (on some measure)

[11] Unconstrained states refer to states where fourth quarter children are not constrained by the law to delay school entry. These would be, for instance, states with January 1^{st} cutoffs. On the other hand, constrained states refer to states with an October 1^{st} cutoff.

from being 6 instead of 5.5 (and lose 4 from being 5 instead of 5.5). For the parents of these children, the benefit of delay does not outweigh the cost of extra child care. All such children enter at age 5. However, 25% of children get a benefit of 12 from entering at age 6 instead of 5.5. The parents of these children all choose to delay the child's entry. The average gain (treatment effect) from delaying entry from age 5.5 to age 6 is 0.75*4+0.25*12 = 6. However, the IV estimate is the average outcome for those born in the second half of the year minus the average outcome for those born in the first half of the year divided by the difference in average age at entry or 0/0.25 = 0. *Even though every child benefits from entering school when older, the IV estimate is that entry age has no effect on the outcome.* As this simple example illustrates, failure to satisfy the monotonicity assumption can produce an estimate with the wrong sign.



Although neither quarter of birth nor legal entry age satisfies monotonicity, it is possible to find an instrument that does. Figure 2 reproduces the last two columns of table 1 for 1952. We can see that first-order stochastic dominance is satisfied: children born in the fourth quarter in states that prohibit them from entering kindergarten in the year that they turn five enter school later than do those in states that permit them to enter. We note that first-order stochastic dominance is only a necessary, not a sufficient, condition for monotonicity.[12]

[12] As discussed in detail in Heckman, Urzua and Vytlacil (2006), an additional variable influencing entry age but not included in the estimation could be correlated with the state law and lead to a violation of monotonicity. The IV estimate would provide an inconsistent estimate of LATE even though the usual

In the next two sections we compare IV estimates of the effect of school entry on educational attainment for different choices of instrument. We use the argument above to propose an instrument that satisfies monotonicity and show that the quarter of birth instrument and the legal entry age instrument give biased estimates of the policy-relevant LATE.

3 Methods: Two Sample Two Stage Least Squares

We estimate the following equation for educational attainment:

$$A_i = \alpha D_i + X_i^{'}\beta + \sum_{j=2}^{4} Q_{ij}\gamma_j + \delta R_i + \epsilon_i \tag{1}$$

where, A_i is the educational attainment of individual i. D_i is the dummy endogenous variable that takes on the value of 1 if individual $i's$ school entry is delayed from the year in which he turns five to the year in which he turns six. Q_{ij} is a set of three dummy variables ($j = 2, 3, 4$) indicating the quarter of birth of the ith individual. X_i is a vector of observable individual characteristics and R_i denotes state controls. Since OLS estimates of α in the above model might be biased by the decision of some parents to accelerate or redshirt their children, we estimate a 2SLS model based on the following first stage equation:

$$D_i = \pi Z_i + X_i^{'}\lambda + \sum_{j=2}^{4} Q_{ij}\theta_j + \varphi R_i + \upsilon_i \tag{2}$$

The binary instrument Z_i equals one if the individual was required by state law to delay kindergarten entry. In other words if the child's month of birth is later than the state kindergarten entry age cutoff date, Z_i equals one and equals zero otherwise. In this setting, LATE implies that we identify the policy relevant parameter i.e. the effect on those individuals who delay enrollment only because they are constrained by the law.

In contrast, it is unclear what the policy relevance of other LATE estimates would be even if they were consistent. For example, suppose that we use legal entry age as an instrument in a country in which everyone enters exactly at the legally permitted. In this case monotonicity is satisfied. Moreover OLS and IV are identical, which simplifies the analysis. The LATE estimator is therefore a least squares approximation of the effect of entering school when one day older. This is a reasonable measure of the effect of making everyone one day older when she enters school. It is therefore a measure of the effect of

requirements for IV are satisfied. For example, if states permitting early entry age also tended to be states with inexpensive childcare, some people who would delay entry and take advantage of the inexpensive childcare in an early entry age state might succeed in entering their child early in a state with a stricter cutoff. If this effect were modest, stochastic dominance could be satisfied even though monotonicity is not.

moving the first day of school one day later if nothing else changed. However, for the most part, moving the first day of school from early September to early October in order to raise the school entry age is not part of the policy discussion. What is under discussion is whether to change the minimum entry age. The LATE estimate using legal entry age may be a very poor estimate of the effect of moving the entry age for a group of students from just under five years old to just under six years old.

Some researchers (e.g. Black, Devereux and Salvanes, 2008) have relied on a regression discontinuity design rather than the instrumental variables approach. At least in our context, taken literally, RD relies on a "no defier" condition rather than the more general monotonicity condition. If the cutoff date for school entry is September 1, then children born on August 31 and who are red-shirted will be only one day older than those born on September 1 who enter when they are first legally permitted to do so. As we further shrink the time range (those born just before and just after midnight), the difference shrinks to zero. So all we require is that any child who would be red-shirted if born just before midnight would comply with the law if born just after midnight. Similarly, any child who would enter before the legal permissible age if born just after midnight would not be red-shirted if born just before. This corresponds to the no defier assumption with a dichotomous treatment and instrument. The RD design captures the effect of being a year older on those affected by the law right at the discontinuity.

Of course, in practice, the discontinuity sample is drawn from a much broader range than a few minutes or even a day. It is common in such cases to include a trend (or separate trends on each side of the discontinuity). Provided the trend is properly specified, this approach still captures the effect right at the discontinuity. It is worth noting that the trend will capture any effect of being a day younger and the changing degree of compliance. However, provided that compliance changes continuously (parents of children born on July 31 do not act in a very different way from those with children born on August 1), this does not affect the interpretation of the coefficient on the discontinuity.

The external validity of the RD estimate is a more serious issue. Policy-makers do not generally consider changing the entry cutoff by an infinitesimal amount. It is much more common to change the date by a few months. If the RD measures the effect of being a year older on children born in late December and who are not red-shirted, it may not provide a good estimate of the LATE for children born in early October and who are not red-shirted since red-shirting is likely to be much more common among the former than among the latter.

This suggests that we might want to look at the fourth quarter/first quarter discontinuity rather than the December 31/January 1 discontinuity. However, if the RD estimate does not control for a trend and the discontinuity sample includes observations that are fairly far from the discontinuity, then the monotonicity assumption may become important again.

To our knowledge, there is no large nationally representative data set with information on school entry age, educational attainment and quarter of birth. To circumvent the lack of data, we use the Two Sample Instrumental Variable (TSIV) procedure developed by Angrist and Krueger (1992,1995). TSIV requires that we have data on the endogenous variable (D_i) and the instrument, Z_i, for a cohort in one data set and the outcome of interest (A_i) and Z_i of the same cohort in another data set. We combine data from the 1960 and 1980 US Census for individuals born in the US between 1949 and 1953. We obtain first stage coefficients from the 1960 Census and use them to predict entry age of the contemporaneous 1980 Census respondents. Instrumental variable estimates are generated by regressing 1980 educational outcomes on the cross-sample fitted value of their entry age. [13]The standard errors are then adjusted to account for the use of a predicted value in the second stage. The appendix gives a detailed description of the method used to consistently estimate the correct asymptotic covariance matrix.

Since we control for quarter of birth (and state), the instrument has a monotonic effect on school entry age. The monotonicity assumption would be violated if there were "defiers" (Imbens and Angrist, 1994; Angrist, Imbens and Rubin, 1996), in other words, if some children born in the fourth quarter enter school early only when they are prohibited from doing so. Although we cannot directly test for such violations, we find them implausible.

Our identification strategy requires that the school entry cutoff date has no effect on the educational outcomes of children who are not constrained by the law. This assumption would be violated if parents do not want their child to be the youngest in class. In this case, they might not redshirt a child born in September in a state with a late cutoff (e.g. January 1), but decide to redshirt in a state with an early cutoff (e.g., October 1).

Table 2 provides some evidence that this "no externality" condition is satisfied. Using the 1960 Census, we show the average entry age by quarter of birth of individuals born between 1949-1953 in two types of states. As this table illustrates, the average school entry age of individuals born in the first three quarters does not vary much by whether they are in a fourth quarter constrained state or not. In other words, raising the minimum entry age does not affect redshirting among those not constrained by the law. In contrast, the distribution of entry age for the fourth quarter differs noticeably between the two types of states.

Moreover, in the 1960s there is significant noncompliance, especially among fourth quarter children, in both types of states. In states with a 10/1 or 9/30 cutoff, almost 45% of

[13] Inoue and Solon (2006) call this the two-sample two-stage least squares (TS2SLS) estimator. They note that in finite samples, the TSIV estimator originally proposed by Angrist and Krueger and the TS2SLS estimator typically used by practitioners are numerically distinct. In addition, they show that the TS2SLS estimator is asymptotically more efficient.

fourth quarter individuals enter school even before they are allowed to enter. On the other hand, in states which allow fourth quarter children to enter early, about 25% redshirt.

Table 2: Distribution of Entry Age (1960 census, 1949-1953 cohorts)

	10/1 or 9/30 cutoff		**1/1 or 12/31 cutoff**	
Birth Quarter				
Born Quarter One	23.97		24.75	
4.5		11.4		10.99
5.5		73.44		72.99
6.5		14.38		14.99
7.5		0.78		1.03
Born Quarter Two	23.84		23.04	
4.25		7.42		7.45
5.25		75.94		73.02
6.25		15.56		18.3
7.25		1.08		1.23
Born Quarter Three	26.92		26.81	
4		7.28		6.76
5		72.97		73.57
6		18.79		18.59
7		0.96		1.08
Born Quarter Four	25.26		25.40	
3.75		4.76		7.64
4.75		40.57		67.21
5.75		47.77		23.54
6.75		6.90		1.62

Note: Constrained to enter at the earliest reasonable age if actual entry age was either too young or too old

4 Data

We use the 1960 US Census Public Use Microdata Sample (PUMS) one percent sample for school entry age data and the 1980 U.S. PUMS five percent sample to measure educational attainment. Both samples have information on quarter of birth.

The main endogenous variable is a dummy variable indicating whether the individual delayed school enrollment from the year he turned 5 to the year he turned 6 or later. Age in quarters was computed as of Census day (April 1, 1960) using information on quarter of birth. The census, however, does not collect school entry age information. School entry age can still be computed using highest grade completed if we assume that no one repeats or skips a grade. We do not know whether children attended kindergarten or entered first grade directly as was common during this period. We treat all individuals as having spent a year in kindergarten. Thus someone who first enrolled in school as a first-grader at exactly the age of 6 would be counted as having entered school at exactly age 5. Based on this assumption, we computed the school starting date for individuals born in the US between 1949 and 1953.

Our identification strategy requires knowledge of exact kindergarten entry cutoff dates for 1954 to 1958, the years in which the individuals in our sample were eligible to enroll in kindergarten. We collected data on state laws regarding kindergarten entry ages. We verified these laws by looking at the US historical state statutes. If the history of the statute indicated a change in the state law in any given year, we examined the state session law to determine the exact form of the change. Children who entered school in states that gave Local Education Authorities the power to set the entry age were deleted from the sample. Table 3 lists the kindergarten entry age cutoff dates for 1958 for the states used in our analysis.

For both samples, we use information on quarter of birth, age, state and cutoff date to determine whether each sample member was born before or after the state cutoff. We delete observations for whom we cannot determine whether the individual was born before or after the cutoff. For example, we drop individuals born in the third quarter in states with a September 1 cutoff. In both data sets, we restrict the sample to individuals whose state of birth and current residence were identical. The sample is restricted to blacks and whites including those of Hispanic origin. For the 1980 sample, we only include individuals who had completed at least one year of schooling. Our final sample includes 96676 observations in the 1960 Census and 373845 observations in the 1980 Census. All regressions include dummies for quarter of birth, sex, race and state and age in quarters and age squared.

Table 3: School Entry Cutoff Dates in 1958

1-Sep	10-Sept/15-Sept	30-Sept/1-Oct	15-Oct/16-Oct	31-Oct/1-Nov	1-Dec	31-Dec/1-Jan	1-Feb
Colorado	Iowa	Alabama	Idaho	DC	California	Connecticut	Pennsylvania
Delaware	Montana	Arkansas	Maine	North Dakota	Illinois	Florida	
Kansas	New Hampshire	Missouri	Nebraska	Oklahama	Louisiana	Kentucky	
Michigan	Ohio	New Jersey		South Dakota	New York	Maryland	
Minnesota	Wyoming	North Carolina		West Virginia	Wisconsin	Mississippi	
Oregon		Virginia				Nevada	
Texas						New Mexico	
Utah						Rhode Island	
						Tennessee	

5 Results

5.1 First Stage

Table 4 presents the first stage results from the 1960 Census for different choices of instrument. Column (1) reports results from the regression of entry age (in years) on one quarter of birth dummy (QOB 1 versus all others). Column (2) uses three quarter of birth dummies (QOB 4 is the omitted quarter). Column (3) shows first stage results using legal entry age as the instrument without quarter of birth controls and finally, column (4) reports estimates from our basic model, controlling for three birth quarters and a binary instrument (delayed by law).[14] Controlling for legally mandated delayed enrollment in column (4), the school entry age monotonically decreases with quarter of birth. Column (4) reveals that individuals born in the first quarter begin school when they are about one-half year older than are those born in the fourth quarter and who are not constrained by state laws. On the other hand, in column (2) the quarter of birth instrument shows a much smaller difference in entry age between the first and the fourth quarter since it fails to control for the more restrictive laws in some states. Note also that the effect of "delayed" is only .37. While some children born in the fourth quarter begin school when they are first allowed to enroll, others are held back an additional year until they are almost 6 years old, and some who are not legally entitled to enroll before age five are nevertheless able to do so.

Table 4: First Stage Estimates: 1960 census (1949-1953 cohorts)

	(1)	(2)	(3)	(4)	Born in 1953*
	Dependent Variable: School Entrance Age				
Born quarter 1	0.2447	0.2685		0.5585	0.4273
	(0.0193)	(0.0374)		(0.0132)	(0.0179)
Born quarter 2		0.1119		0.4024	0.5988
		(0.0377)		(0.0128)	(0.0161)
Born quarter 3		-0.0834		0.1942	0.4468
		(0.0392)		(0.0118)	(0.0161)
Delayed by Law				0.3664	0.1985
				(0.0397)	(0.0161)
Legal entry age			0.4141		
			(0.0551)		
Observations	96676	96676	96676	96676	19949
R-squared	0.09	0.11	0.09	0.11	0.18

Note: Robust standard errors clustered by state/quarter of birth.
Controls: state fixed effects, age in quarters, age square, race (white/black) and sex.
Sample restricted to individuals for whom state of birth is identical to birthplace.

One concern with the entry age variable is that since we assume there is no grade retention, we are overestimating the entry age. This is especially problematic since past

[14] Note that this specification is isomorphic to one in which legal age is used as the IV and quarter of birth is included in the structural equation. This specification can be found in the literature as a robustness check (Elder and Lubotsky, 2006).

research has shown that the probability of repeating a grade is related to school entry age. Although we do not have information on grade retention in the Census, we can minimize the error in measuring the entry age variable by restricting the sample to the youngest cohort. The fifth column of Table 4 restricts the sample to those born in 1953. If one assumes that entry patterns were constant from 1949 to 1953, then the difference between the baseline estimates in column (4) and those obtained using only the 1953 data reflect the effect of grade retention. In this case, estimates based on 1953 data would be preferred. Estimates using the 1953 only first-stage can be obtained by multiplying coefficient on "delayed" in the baseline model by .3664/.4273 or .8575.

It is also worth noting that, using the 1953 data, the difference in entry age between those born in the second and third quarter is almost exactly .25, suggesting that monotonicity would apply to a sample of individuals born in these quarters. This, in turn, would mean that it is possible to compute a LATE based on these samples. However, it is not clear that this LATE would be of any policy interest.

Table 5: Reduced Form Estimates 1980 census (1949-1953 cohorts)

	QOB	Legal Age	Delayed
Dependent Variable: Educational Attainment			
Legal entry age/Delayed		-0.0290	-0.0029
		(0.0211)	(0.0265)
Born in quarter 1	-0.0444		-0.0689
	(0.0081)		(0.0222)
Born in quarter 2			-0.0459
			(0.022)
Born in quarter 3			-0.0122
			(0.0234)
Observations	373845	373845	373845
R-squared	0.05	0.05	0.05

Note: Robust standard errors clustered by state/quarter of birth.
Additional controls for state, age in quarters, age squared, race (white/black) and sex.

5.2 Reduced-Form and TS2SLS Estimates

Table 5 reports reduced-form estimates from the 1980 Census. In column (1), which gives the reduced form when the instrument is "born in first quarter," the instrument is associated with a large negative effect on educational attainment. In column (2), legal entry age instrument shows a somewhat smaller and statistically insignificant adverse effect. Finally, the last column indicates that controlling for quarter of birth, there is almost no effect of

Table 6: Two Sample Instrumental Variable Estimates 1960-1980 census

	QOB	Legal Age	Delayed
	Dependent Variable: Educational Attainment		
Predicted Entrance Age	-0.1815	-0.0700	-0.0078
	(0.0422)	(0.0450)	(0.0727)
Born in quarter 1			-0.0645
			(0.0229)
Born in quarter 2			-0.0427
			(0.0146)
Born in quarter 3			-0.0107
			(0.0152)
Observations	373845	373845	373845
R-squared	0.05	0.05	0.05

Note: Moulton-corrected standard errors in parentheses.
Additional controls for state, age in quarters, age squared, race (white/black) and sex.

delayed school entry on educational attainment.

Table 6 combines estimates from the 1960 and 1980 Censuses. Using first stage coefficients reported in Table 4, we predict entry age for the 1980 Census respondents. TS2SLS estimates are generated by a regression of 1980 educational outcomes on the predicted entry age. Using the method described in the appendix, we correct the standard errors to account for the fact that the predicted value of school entry age is used in the second stage. In addition, the standard errors are adjusted for clustering (at the level of state*quarter of birth) using a parametric Moulton (1986) correction factor.[15]

When we use "born in the first quarter" as our instrument, consistent with Angrist and Krueger, we find a large negative effect of school entry age on educational attainment. When we use legal entry age (not controlling for quarter of birth), we find a smaller but still substantial adverse effect that falls short of statistical significance at conventional levels and is therefore consistent with the zero effect in Bedard and Dhuey. Finally, when we use the consistent estimator that meets the monotonicity requirement, our estimate is very close to zero.

We also study the effect of delayed enrollment on other measures of educational attainment namely, high school Dropout/completion and college attendance. We have also

[15] Stata does not have a command to compute the Moulton standard errors. We use the user written command provided by Joshua Angrist and Steve Pischke.

Table 7: **TS2SLS Estimates: Blacks Only**

	Entire Sample	Blacks Only
Attainment	-0.0078	0.2759
	(0.0727)	(0.2250)
High School Dropout	-0.0136	-0.0832
	(0.0107)	(0.0425)

Note: Moulton-corrected standard errors in parentheses.

looked at the differences in outcomes by sex, race and race and sex interacted, but do not find any statistically significant effect. We do not find any effect for whites or for either sex separately. However, as shown in Table 7, we find a nontrivial and marginal statistically significant effect of delay on the dropout rate among blacks. For blacks delaying entry to school is associated with a decline in the dropout rate of about 8 percentage points. Our point estimates also suggest that delayed entry increases educational attainment among blacks by a nontrivial quarter of a year. However, the coefficient is not significant at conventional levels. We do not want to put too much weight on this finding. After all, we have looked for significant effects on several overlapping groups using multiple measures of educational attainment. Finding a t-statistic of just under 1.96 in one specification for one group is not all that unlikely. However, it is plausible that blacks were affected more adversely by delay than were other groups. For blacks, it is particularly important to note the historical nature of the finding since we are looking at students starting school during a period that preceded the 1964 Civil Rights Act and when blacks were hugely disadvantaged both in terms of school quality and parental income. It is plausible that black children (and other children from disadvantaged backgrounds) who enter school early did so for financial reasons and were frequently pushed ahead before they were sufficiently mature.

6 Optimal School Entry Age

The three estimates in Table 6 may differ for one or both of two reasons. First, if the failure of the monotonicity assumption is important, quarter of birth and legal entry age instruments do not provide a consistent estimate of a local average treatment effect. Second, the local average treatment effects captured by the instruments may differ. To determine the importance of these two sources of divergence, we propose and estimate a model of optimal school entry age. We then use the model to examine the relations between the two standard IV estimators and the effects they are intended to measure.

6.1 Model

Every child has an optimal school entry age, E_i^* , where

$$E_i^* = a_0 + a_1 \widetilde{E}_i$$

We assume that the random component, $\widetilde{E}_i$ is distributed $Beta(\alpha, \beta)$ with the two shape parameters α and β. The parameters a_0 and a_1 determine the bounds of the optimal entry age distribution, a_0 gives the lower bound while $a_0 + a_1$ sets the upper bound. We allow a_0 to depend on the state entry age law. However, because we have data on quarter of birth (as opposed to month of birth), we restrict the analysis to two types of states. The unconstrained states (u) refers to states with a either a 1/1 cutoff or a 12/31 cutoff. The second type of state, the fourth quarter constrained state (c), is restricted to states with 9/30 or 10/1 cutoff.

We introduce a shift parameter for being in a constrained state:

$$a_o^c = a_o^u + \lambda$$

This implies that raising the minimum entry age for fourth quarter children may affect the optimal entry age for everyone else. By allowing the optimal entry age to be affected by school entry age laws, we are allowing for spillover effect of laws. Existence of such externalities would be a violation of the exclusion restriction required for identification using instrumental variables, including our own, based on legal entry age.

Let E_i be the actual age at which a child begins school. E_i would differ across children because of differences in quarter of birth and school cutoff. We assume that students suffer an education penalty if they enter at an age other than their optimal entry age E_i^*. For example, a student who is born on March 1 and whose optimal entry age would be age 5 (if school started on March 1), is now forced to enter at age 5.5 because school begins on September 1. She suffers a loss associated with being six months away from her optimal entry age. We assume that the education loss is quadratic in the absolute departure from optimal entry age. Thus, ultimate educational attainment is given by:

$$S_i = S_i^* + \mu_1 * |E_i - E_i^*| + \mu_2 * (E_i - E_i^*)^2$$

S_i^*, which is unobserved, is the educational attainment the individual would have attained if she had entered at exactly her optimal entry age. We assume that S_i^* is independent of quarter of birth and state cutoff date. This assumption rules out season of birth effects (Bound and Jaeger, 1995, 2000; Buckles and Hungerman, 2009) and also allows us to focus attention on the quadratic cost term.

Our choice of this particular form is driven by the paucity of data. As discussed below, we use the data to identify six parameters. We impose that $a_o^u + a_1$ equals 7. In other words the highest optimal school entry age in an unconstrained state is seven years old. We choose this restriction because all states require children to enter school by the time they are eight years old. Since in most states, kindergarten is not required, eight year olds starting school would typically enter first grade. This is equivalent to requiring children to begin kindergarten at age seven in our model.

6.2 Indirect Inference

We use indirect inference to estimate the six parameters of the model ($a_0, \alpha, \beta, \lambda$, and μ_1 and μ_2), with $a_1 = 7 - a_0$, so that the moments from the simulation match the moments from the data. We generate 10,000 draws from the beta distribution.

For simplicity, we assume that children born in quarter 1 are born on 2/15, quarter 2 on 5/15, quarter 3 on 8/15, and quarter 4 on 11/15. Further we assume that the first day of school each year is August 15th in every state. This implies that Quarter 1 students can enter school at age 4.5, 5.5, or 6.5. Similarly, those born in quarter 2 can enter at 4.25, 5.25 or 6.25 and so on for the third and the fourth quarter.

We do not impose that individuals enter school at the date that is closest to their optimal entry age. Instead we assume that individuals with the lowest optimal starting age are the ones among those born in a given quarter who enter when youngest. In other words, if we observe in the data that 10% of first quarter children enter at age 4.5, we assume that these are the 10% of the first quarter children with the lowest optimal entry age.

Based on these assumptions, we use the distribution of entry age (1949-1953 cohorts) from the 1960 Census to generate simulated data. Thus, we allocate individuals to their entry age in the simulated data consistent with their quarter of birth and whether they live in a 4th quarter constrained state or not. Table 3 shows that there are 15 states in the sample with cutoff dates corresponding to the c and u states. The Census distribution of entry age that we use to generate the simulated data has been shown earlier in Table 2. As previously noted, Table 2 does not suggest a spillover effect of increasing entry age for those born in the fourth quarter on those born in the first three quarters. We have also noted that in the 1950s the laws were not strictly enforced as there is a lot of noncompliance in this sample.

Next, we regress educational attainment from the 1980 census on three quarter of birth dummies, age in quarters and its square and state dummies, separately for the two types

Table 8: Educational Attainment and Quarter of Birth by State Type

	Unconstrained States				Constrained States			
	Estimated	Simulated			Estimated	Simulated		
		I	II	III		I	II	III
Quarter 1	-0.076 (0.013)	-0.076	-0.079	-0.078	-0.093 (0.010)	-0.093	-0.095	-0.096
Quarter 2	-0.019 (0.016)	-0.019	-0.019	-0.018	-0.046 (0.014)	-0.046	-0.038	-0.038
Quarter 3	0.023 (0.012)	0.023	0.018	0.017	-0.004 (0.007)	-0.004	-0.004	-0.004

Note: Robust SE clustered by state/quarter of birth. Controls include state dummies, age in quarters and its square.
N=292771

of states to get the vector of coefficients $\widehat{\beta}_{data}$(i.e. a total of six moments, coefficients on three quarter of birth dummies in each type of state). These coefficients are the difference in average education between those born in each of the first three quarters and those born in the fourth quarter in each type of state. Identification in this model depends only on within state-type education differences since we are not using the difference in average educational attainment between the two types of states.

Finally, we characterize the loss function as the sum of the squared deviations between the regression coefficients from the simulated data and the actual regression coefficients weighted by the inverse of the variance-covariance matrix of the estimates, $\widehat{\Sigma}$.

More formally, the objective of our indirect inference simulations is to choose parameters of optimal entry age distribution (α, β, a_0, plus the shift parameter for constrained states, λ) and of the education loss function (μ_1 and μ_2) to minimize the following loss function:

$$(\widehat{\beta}_{data} - \widehat{\beta}_{sim})'\widehat{\Sigma}^{-1}(\widehat{\beta}_{data} - \widehat{\beta}_{sim})$$

6.3 Simulation Results

Table 8 shows results from regressions of educational attainment on three quarter of birth dummies using the actual and simulated data. It shows the average difference in educational attainment between the fourth quarter and the three other quarters. Consistent with Angrist and Krueger (1991), in both types of states individuals born in the first quarter get less education than do those born in the other quarters.

Table 8 also shows the estimates based on simulated data from the three variants of the model. Model I is the unconstrained model previously described. It fits the actual data almost exactly.[16] Model II constrains the quadratic term μ_2 to be zero, while model

[16] Although the literature on indirect inference assumes that if the number of model parameters equals the

III constrains both μ_2 and λ, the shift parameter between constrained and unconstrained states both to be zero. It is evident that neither of these restrictions prevents the model from fitting the empirical parameters quite precisely. We address this point more formally in the next table.

The top panel of table 9 shows the fitted parameters. The first column gives the parameters for the base model. For the most part, the estimated parameters are plausible. The lowest optimal entry age is just under four years old; the mean is about 5.3, and the distribution is somewhat skewed so that the median is lower. There is little evidence of an external effect from raising the minimum school entry age. The distribution of optimal entry ages is less than .01 higher in states with an October 1 cutoff than in states with a cutoff at the end of the year. The main problem with the model in the first column is that the cost parameters are implausible. The estimates imply that individuals benefit from deviating from their optimal entry age until the deviation exceeds six months and that those who deviate by a year do almost as well as those who enter at exactly their optimal age.

Table 9: Model Parameter Estimates and Implied Treatment Effects

	Baseline	No Quadratic Term	No Quadratic, No λ
Model Estimates			
Beta Distribution			
α	1.64	0.83	0.81
β	2.21	2.29	2.25
Bounds of optimal age			
a_0	3.97	4.49	4.5
a_1	3.03	2.51	2.5
Loss from Deviating			
μ_1	6.89	-0.67	-0.67
μ_2	-6.92	0*	0*
External Effect			
λ	0.01	0.01	0*
Loss Function	0.00	0.78	0.81

*Coefficients constrained

a_0 and a_1 are constrained to sum to 7

We therefore restrict the quadratic term to be zero. The results are shown in the second column. We note first that the restriction cannot be rejected. The loss function only increases to .78, well below the critical value for a chi-squared with one degree of freedom.[17] The lowest optimal entry age is now estimated to be four and a half, suggesting that some

number of empirical parameters, the fit must be perfect, it is easy to show that this need not be the case even when the underlying model is correctly specified.

[17] We also experimented with restricting the linear term to be zero, on the grounds that the derivative of education with respect to entry age should be smooth around the optimal entry age. While we cannot reject this model at conventional levels of significance, the loss function with this specification is substantially higher than with the restriction on the quadratic term.

children born in the first quarter should enter in the year they turn four, something that was relatively unusual even in the 1950s. The parameters of the beta-distribution imply that optimal entry age is very skewed. The mean is 6.3 but the median is only 5.0. Most children would benefit from entering when relatively young, but some would be better off being significantly older than the norm.

We continue to find no evidence that constraining the age at which children born in the fourth quarter can enter school has any effect on the optimal age for other children. The estimated value of λ is less than .01. Therefore in the third column of Table 9, we restrict the value of λ to be zero.

Given the very low value of λ, it is not surprising that this restriction cannot be rejected. The loss function increases by only .03, and the remaining results do not change noticeably.

If we accept the two parameters restrictions imposed in the third column, then, in principle, it is possible to unconstrain a_1 which determines the oldest age at which it is optimal for any child to entry school. It will be apparent that since the loss function is only .81, we will not be able to reject that the $a_0 + a_1$ equals 7. In practice, a_1 is very imprecisely estimated, and we have been unable to converge the model without this restriction.

Both the second and third columns of Table 9 imply that the cost of entering at the wrong age is large. A child who enters at exactly age 5 and who should have entered at exactly age 6 or vice versa loses, on average, about two-thirds of a year of education. Of course, almost everyone can choose to enter within six months of her optimal entry age. The major exceptions are individuals with very young optimal entry ages who are born in the first quarter or who are born in the fourth quarter in states requiring them to wait until the year in which they turn six to enter school.

6.4 Reconsidering the Instruments

As discussed earlier, there are at least two reasons that the IV estimates using (first) quarter of birth, legal age, and delayed by law may diverge. The first is that the failure of the monotonicity assumption makes either one or both of the first two estimators inconsistent. The second is that they measure different local average treatment effects.

Table 10 shows the results of applying each of the IV estimators to the data generated by our model. We calculated the estimates for the models in columns 2 and 3 of Table 9. Since they gave identical results to two decimal places, we do not distinguish between them here.

The first column of Table 10 repeats the results from Table 6. The corresponding rows in the second column show the estimates applied to our data. Although our parameters were not chosen to match the three IV estimates, the model fits the broad pattern found in the data. The "born in first quarter" instrument shows the most adverse effect of delaying entry while the "delayed by law" instrument finds the least adverse and possibly positive effect. In each case, the estimate derived from the model lies within the confidence interval of the actual estimate.

Table 10: Effects of Entry Age on Education

	From Data	From Model
IV - Quarter 1	-0.18 (0.04)	-0.23
True LATE - Quarter 1		-0.26
IV - Legal Age	-0.07 (0.05)	-0.09
True LATE - Legal Age		-0.22
True LATE - Delayed	-0.01 (0.07)	0.06

Next we ask how well each IV estimator would capture its intended LATE if the true world were generated by our model. What LATE should each estimator capture? In the absence of monotonicity the concept is not well-defined. For the quarter-of-birth instrument, our solution is to treat all induced entry age changes as positive. Thus, we define LATE as the local average treatment effect of taking someone born in the first quarter and having him be born in each of the other three quarters (in each of the two types of states) but accounting for the sign of the effect on entry age. To do this, we rank every observation in increasing order of the optimal entry age in each of the four quarters. We match the corresponding lowest optimal entry age in the first quarter with the lowest in the second, third and fourth quarters. Similarly, the second lowest in the first quarter is matched with second lowest in each of the other quarters and so on. Thus, we match up individuals in the first quarter with individuals with the same ranking in the optimal entry age distribution in the other three quarters. Next, we calculate the difference in entry age and the difference in education for each of the observations to get a total of three differences for each observation in each type of state and six differences overall. Since we want to satisfy monotonicity, if the difference in entry age is negative, we flip the sign of the entry age and education attainment differences. Using these numbers, we calculate the true LATE as the total loss in education divided by the difference in average entry age.

The true LATE defined in this way is given in the second row of the last column of Table 10. At least in the world represented by our model, the IV estimator is somewhat biased but not badly so. It is off by about .03.

Although the legal entry LATE relies on variation in both birth date and state laws, it seems to us that the goal is to estimate the effect of a small increase in entry age (from being born on, for example, February 1 rather than February 2) rather than some strange combination of small increases due to birth dates and large increases due to state law. We therefore calculate the (numeric) derivative of educational attainment with respect to an increase in entry age for all individuals in our sample and take the average. The result of this exercise is shown in the fourth line of the last column in Table 10. It is evident that if this is the LATE that "legal age" is intended to capture, then it badly fails to do so. The estimated LATE is quite far from the true LATE.

By construction, using our approach, we get a consistent LATE estimate of the effect of the policy change of moving from a December 31 to an earlier cutoff on the educational attainment of children born in the fourth quarter (i.e. those children whose behavior is affected by the law). However, it is important to recognize that our estimates assume that there are no externalities from this change. We find no evidence of the existence of such externalities, but this is quite different from finding strong evidence of their absence. Conditional on this caveat, those children whose entry is delayed, on average, are not harmed and may benefit slightly from the delay.

6.5 Policy Experiments

An important policy question that arises from our analysis is whether our results would hold in the current school system where school entry laws are relatively strictly enforced. The weakly enforced cutoff dates in the 1950's may not be applicable to the debates involving school entry age today. Schools today are under great pressure to adhere to strict standards. As discussed in the introduction, a variety of factors have pushed states and districts to increase entry age requirements and enforce them more strictly, but it is very uncertain as to whether such policies are beneficial.

To study the effect of delaying school entry on attainment in recent years, we use the simulated data to perform some policy experiments. First, we look at the effect of moving from a January 1 cutoff to an October 1 cutoff around the 1950's, a period when such cutoffs were very loosely enforced. Second, we consider what would have happened had their been a strict October 1 cutoff.

Table 11 reports the results from these two experiments. In the first column, we explore the effect of the policy in a model that permits the policy to affect the optimal entry age. In this experiment we assume that the small differences we observe between the entry age decisions of parents of children born in the first three quarters and living in constrained states and those of their counterparts in unconstrained states reflect responses to the change

Table 11: Effect of Raising Entry Age by Type of Law Enforcement: Assorted Models

	No Quadratic Term	No Quadratic, No Externality*
	Weak Enforcement	
Change in Educational Attainment		
All	0.008	
Q1	0.006	
Q2	0.003	
Q3	0.004	
Q4	0.022	0.022
Percent Increasing Educational Attainment* *		
All	61.8	
Q1	79.2	
Q2	70.5	
Q3	50.1	
Q4	47.6	56.3
	Strong Enforcement	
Change in Educational Attainment		
All	-0.060	
Q1	0.005	
Q2	0.003	
Q3	0.000	
Q4	-0.246	-0.250
Percent Increasing Educational Attainment* *		
All	54.6	
Q1	77.9	
Q2	70.4	
Q3	50.0	
Q4	20.4	18.4

*With no externality, there is no effect on other quarters or on those who do not change their behavior.
**Of those changing their educational levels

in optimal entry age. When we use the model that assumes no externalities, we assume such entry age changes are random and ignore them.

The top panel shows the effect of weak enforcement. In both models there is a slight increase in average education, with the benefits accruing primarily among children born in the fourth quarter. Strikingly, even though the average child born in the fourth quarter benefits, in each model roughly half of those whom the law causes to delay entry benefit and half do not. When we allow the policy to increase the optimal entry age for everyone, most children born in the other quarters benefit, but the average gains are very small.

The lower panel shows the results from the policy experiment with strict enforcement. In the first column, we assume that the increase in optimal entry age is proportional to the increase in average entry age, so there is a bigger externality when enforcement is strict. Both with and without an effect on optimal entry age, we find that moving from a January 1 cutoff lowers average educational attainment by about .06 years, with a large adverse effect on those born in the fourth quarter and positive effects on those born in the first two quarters in the scenario allowing for externalities.

When laws were weakly enforced, the constrained children (those born in the fourth quarter) had the option to enter school earlier than officially permissible. We see ample evidence of this happening in our data. In this environment, overall, children benefited, in terms of higher educational attainment, by moving to an October 1st cutoff. However, the policy experiment suggests that, in an environment where laws are strictly enforced, constraining fourth quarter children to enter late hurts these children and reduces average educational attainment.

7 Conclusion

In this paper, we argue that previous studies that have used IV to deal with the endogeneity of school entry age have focused on a LATE of no real policy or practical interest. Moreover, they have failed to provide consistent estimates of the LATE because of the failure of the monotonicity assumption. As a practical matter, this turns out to be a bigger problem for the "legal age" instrument than for the "first-quarter birth" instrument.

Our two sample results suggest that the quarter of birth instrument yields severely biased estimates of the policy-relevant LATE. The born in first quarter instrument, consistent with Angrist and Krueger, gives a large negative effect of school entry age on educational attainment. When we use legal entry age (not controlling for quarter of birth), we find a smaller adverse effect but one that falls short of statistical significance at conventional levels (consistent with the zero effect in Bedard and Dhuey). We propose an instrument that satisfies the monotonicity assumption and gives a consistent estimates of the policy-relevant LATE: the effect of requiring a child to enter school in the year she turns six when she would otherwise have entered a year earlier. The results are consistent with no important policy effect *as the policy was practiced in the 1950s.*

However, over the last fifty years, school entry age laws have become noticeably stricter both in requiring children to be older before entering school and through stricter enforcement of the laws limiting entry although they generally continue to permit redshirting. We find that stricter enforcement of the laws in the 1950s would have had adverse effects on educational attainment. While we do not know whether the results continue to apply today, they do provide evidence of considerable variation in optimal entry age and therefore suggest that having a waiver policy that gives constrained children the choice to enter earlier than the legally established age could increase educational attainment, particularly among groups that have high dropout rates.

Appendix: Standard Error Derivation

Let the first stage be

$$Y_{ic} = X_{ic}B_1 + cD_{ic} + \alpha_c + \varepsilon_{ic} \tag{A1}$$

where observations are indexed by i and grouped in clusters indexed by c. D is the excluded instrument. Within each cluster c, the $(Y_i, X_i)'s$ are correlated, but (Y_i, X_i) from different clusters are independent. Let α_c be the random component specific to cluster c and ε_{ic} is the individual specific error term.

For convenience we can write the first stage as

$$Y_{ic} = Z_{ic}\Gamma + \alpha_c + \varepsilon_{ic}$$

Let the structural equation be

$$\begin{aligned} y_{ic} &= X_{ic}B_2 + \gamma Y_{ic} + \mu_c + \nu_{ic} \qquad \text{(A2)} \\ &= X_{ic}B_2 + \gamma(Z_{ic}\Gamma + \alpha_c + \varepsilon_{ic}) + \mu_c + \nu_{ic} \\ &= X_{ic}B_2 + \gamma Z_{ic}\Gamma + \delta_c + \zeta_{ic} \\ &= X_{ic}B_2 + \gamma Z_{ic}\widehat{\Gamma} + \delta_c + \zeta_{ic} + \gamma(Z_{ic}(\Gamma - \widehat{\Gamma}) \end{aligned}$$

Let $X^* = [X \;\; Z\widehat{\Gamma}]$ and $B = [B_2 \;\gamma]$.

Then

$$V(\widehat{B}) = E(X^{*\prime}X^*)^{-1}X^{*\prime}\varpi\varpi' X^*(X^{*\prime}X^*)^{-1}$$

where ϖ is the error term defined above i.e.

$$\varpi = \delta_c + \zeta_{ic} + \left[\gamma Z_{ic}(\Gamma - \widehat{\Gamma})\right]$$

Each of the error terms is orthogonal to Z. Therefore the TS2SLS covariance matrix in the presence of clustering is given by:

$$V(\widehat{B}^{TS2SLS,Moulton}) = (X^{*\prime}X^*)^{-1}X^{*\prime}\Omega X^*(X^{*\prime}X^*)^{-1} + \gamma^2(X^*X^*)^{-1}X^{*\prime}ZV(\widehat{\Gamma})Z'X^*(X^{*\prime}X^*)^{-1} \tag{A3}$$

where Ω is a block diagonal matrix with diagonal elements ω_c (the intra-cluster correlation matrix for each cluster c)

$$\omega = \begin{bmatrix} \sigma_\delta^2 + \sigma_\zeta^2 & \sigma_\delta^2 & . & . & \sigma_\delta^2 \\ \sigma_\delta^2 & \sigma_\delta^2 + \sigma_\zeta^2 & & & \\ . & & . & & \\ . & & & . & \\ \sigma_\delta^2 & & & & \sigma_\delta^2 + \sigma_\zeta^2 \end{bmatrix} \tag{A4}$$

and

$$V(\widehat{\Gamma}) = (Z'Z)^{-1}Z'\Omega_2 Z(Z'Z)^{-1}. \tag{A5}$$

It is easy to show that this formula reduces to the asymptotic covariance matrix formula for TS2SLS estimator derived by Inoue and Solon (2008). However, we also correct for the possibility of Moulton clustering in each stage.

References

[1] Allen, Jeremiah and Roger Barnsley, "Streams and Tiers: The Interaction of Ability, Maturity, and Training in Systems with Age-Dependent Recursive Selection," *Journal of Human Resources* 28 (Summer 1993): 649-659.

[2] Angrist, Joshua D. and Guido W. Imbens, "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90 (June 1995): 431-442.

[3] Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association* 91 (June 1996): 444-472.

[4] Angrist, Joshua D. and Alan B. Krueger, "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106 (November 1991): 979-1014.

[5] Angrist, Joshua D. and Alan B. Krueger, "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 (June 1992): 328-336.

[6] Barua, Rashmi and Kevin Lang, "Kindergarten Entry Age and Academic Performance" unpublished, 2008

[7] Bedard, Kelly and Elizabeth Dhuey, "Is September Better than January? The Effect of School Entry Age Laws on Skill Accumulation" unpublished, 2008

[8] Bedard, Kelly and Elizabeth Dhuey, "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *Quarterly Journal of Economics* 121 (November 2006): 1437-1472.

[9] Black, Sandra, Paul Devereux and Kjell Salvanes, "Too Young to Leave the Nest? The Effects of School Starting Age," NBER Working Paper #13969, April 2008.

[10] Bound, John, and David A. Jaeger, "Do Compulsory School Attendance Laws Alone Explain the Association between Quarter of Births and Earnings?" *Research in Labor Economics,* 19 (2000): 83-108.

[11] Bound, John, David A. Jaeger and Regina M. Baker, "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endoge-

nous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90 (June 1995): 443–450.

[12] Buckles, Kasey and Dan Hungerman, "Season of Birth and Later Outcomes: Old Questions, New Answers," NBER Working Paper # 14573, September 2008.

[13] Cahan, Sorel and Nora Cohen, "Age versus Schooling Effects on Intelligence Development," *Child Development*, 60 (October 1989): 1239-1249.

[14] Cascio, Elizabeth U. and Ethan G. Lewis, "Schooling and AFQT: Evidence from School Entry Laws," *Journal of Human Resources*, 41 (Spring 2006): 294-318.

[15] Datar, Ashlesha, "Does Delaying Kindergarten Entrance Give Children a Head Start?" *Economics of Education Review*, 25 (February 2006): 43-62.

[16] Deming, David and Susan Dynarski, "The Lengthening of Childhood." *Journal of Economic Perspectives*, 22(3): 71–92. (Summer 2008)

[17] Dhuey, Elizabeth and Stephen Lipscomb, "What Makes a Leader? Relative Age and High School Leadership," *Economics of Education Review*, 27 (April 2008): 173-183.

[18] Dobkin, Carlos and Fernando Ferreira, "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" unpublished, 2007.

[19] Elder, Todd E. and Darren H. Lubotsky, "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, forthcoming.

[20] Fertig, Angela and Kluve, Jochen, "The Effect of Age at School Entry on Educational Attainment in Germany" IZA Discussion Paper 1507, 2005.

[21] Fredricksson, Peter and Björn Öckert, "Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance" IFAU Working Paper 2006:12, 2006.

[22] Heckman, J., S. Urzua, and E. Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, August 2006, Vol. 88, No. 3, Pages 389-432

[23] Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (March 1994): 467-475.

[24] Inoue, Atsushi and Gary Solon, "Two-Sample Instrumental Variables Estimators," *Review of Economics and Statistics*, forthcoming.

[25] Lang, Kevin, “Ability Bias, Discount Rate Bias and the Return to Education,” unpublished, 1993.

[26] Mayer, Susan E. and David Knutson, “Does the Timing of School Affect How Much Children Learn?” in Mayer, S. E. and P. E. Peterson, eds., *Earning and Learning: How School Matters*, Brookings Institution and Russell Sage Foundation: Washington, DC, 1999, 70-102.

[27] McCrary, Justin, and Heather Royer, “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth,” NBER Working Paper No. 12329, 2006.

[28] McEwan Patrick J. and Joseph S. Shapiro, “The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates” *Journal of Human Resources,* 43 (Winter 2008): 1–29.

[29] Moulton, Brent R, "Random group effects and the precision of regression estimates," *Journal of Econometrics*, Elsevier, vol. 32(3), pages 385-397 (August 1986)

[30] Puhani, Patrick A. and Andrea M. Weber, “Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age at School Entry in Germany,” *Empirical Economics*, 32 (May 2007): 359-386.

[31] Stipek, Deborah., "No Child Left Behind comes to preschool". *Elementary School Journal* (2006), 106(5),455-465

[32] West, Jerry, Anne Meek and David Hurst, “Children Who Enter Kindergarten Late or Repeat Kindergarten: Their Characteristics and Later School Performance,” *Education Statistics Quarterly*, 2 (Fall 2000): 21-24.

[33] Yitzhaki, Shlomo, “On Using Linear Regressions in Welfare Economics,” Hebrew University of Jerusalem Department of Economics Working Paper No. 217, 1989.