A Note on Why Quarter of Birth is Not a Valid Instrument for Educational Attainment

Dionissi Aliprantis

University of Pennsylvania

5. October 2007

A NOTE ON WHY QUARTER OF BIRTH IS NOT A VALID INSTRUMENT FOR EDUCATIONAL ATTAINMENT

DIONISSI ALIPRANTIS

1. Angrist and Krueger (1991)

The goal in Angrist and Krueger (1991) is to use variation in schooling attainment unrelated to ability to identify the effect of attending an additional year of school. The paper uses date of birth and compulsory education laws as a natural experiment to assign one group of students an extra year of education. While the paper rightly focuses much attention on demonstrating the effectiveness of compulsory education laws, it does not give sufficient consideration to the problems generated for the IV framework if there is significant redshirting[1] around entrance cutoff dates. This problem is distinct from those discussed in Bound and Jaeger (2000), which assumes "the typical compulsory school attendance law requires students to start first grade in the autumn of the calendar year in which they turn 6," precluding the possibility of redshirting.

1.1. **Monotonicity.**

The implicit assumption needed to identify the effect of attending an additional year of school is that children enroll in school when first eligible. Without this implicit assumption, the explicit assumption of Monotonicity (Angrist and Imbens (1995)) does not hold. As reported in Aliprantis (2007), in 1998 about 30% of parents in the US chose to delay their children's enrollment in primary school if their child's birthday was within two months of their school's entrance cutoff date. This rate was almost one in four for the youngest quarter of eligibility to enroll. Furthermore, this selection was not random, as will be discussed in Section 2.

To illustrate the problem with this implicit assumption suppose, for the sake of simplicity, that all children born in quarter 1 (Q_1) are born on February 15th, quarter 2 (Q_2) on May 15th, quarter 3 (Q_3) on August 15th, and quarter 4 (Q_4) on November 15th. Assume that school starts on August 15th and the cutoff date for all children to turn six is December 31st of the academic school year. Then Table 1 shows the pattern of ages at entry of first grade we will observe.

Now assume that all children must attend school until they reach 16 years of age and that we are looking only at those children who would like to drop out once they turn 15 years old. Then Table 2 shows the years of completed schooling at which students are eligible to drop out conditional on redshirting and quarter of birth. The spirit of the identification scheme used in Angrist and Krueger (1991) is that since laws apply to a student's age, and not their years of completed schooling,

Date: October 5, 2007.

Email: jalipran@econ.upenn.edu

I acknowledge financial support from the Institute of Education Statistics.

[1]Redshirting is the term given to parents voluntarily delaying their children's enrollment

variation in students' birth dates means they face different laws regarding the years of schooling they must complete. In our example, students born in Q_4 must stay in school until they complete 10.25 years of schooling, while individuals born in Q_1 only have to complete 9.5 years of schooling before being eligible to drop out. Thus comparing these groups will tell us the treatment effect (ie, the effect of attending an extra year of school against one's will).

There is a problem with this scheme if there is significant redshirting that is not random with respect to expected educational attainment. Let us begin with the framework in Angrist and Imbens (1995), comparing only those children born in the first and last quarters of the year. Define treatment as being eligible to enroll when young (ie, being born in Q_4). Let S_1 be the years completed for those exposed to treatment, and S_0 the years of completed schooling for those not exposed to treatment. Then if all children enroll when eligible and comply with the compulsory education law, we have from Table 2 that $S_1 = 10.25$ and $S_0 = 9.5$. Thus

$$E[Y_i|Q_4] - E[Y_i|Q_1] = E[Y_i|S_1] - E[Y_i|S_0] = E[Y_i|S_0 + 0.75] - E[Y_i|S_0].$$

However, consider what happens if not all parents enroll their child when first eligible. Suppose that there are two types, H and L. Assume that both types enroll when eligible when born in Q_1, but that only those of type H enroll when eligible if they are born in Q_4. Let B_j be the cohort of children born j months before their school's cutoff date. Table 3 shows that this assumption is reasonable in the ECLS-K data set. For those of type H, $S_1 = S_0 + 0.75$. However, for those of type L, $S_1 = S_0 - 0.25$. Thus for those of type H, $S_1 - S_0 = 0.75 > 0$, but for those of type L, $S_1 - S_0 = -0.25 < 0$. This violates the assumption of Monotonicity.

To illustrate why this creates a problem, consider what would happen if those who would drop out when 15 years of age are all of type L. In this case we would expect that those entering school when younger tend to complete more schooling - precisely what is reported in Dobkin and Ferreira (2007). Then comparing outcomes of groups Q_4 and Q_1 would yield something very different from the parameter of interest:

$$E[Y_i|Q_4] - E[Y_i|Q_1] = E[Y_i|S_1] - E[Y_i|S_0] = E[Y_i|S_0 - 0.25] - E[Y_i|S_0].$$

This leaves us unable to interpret the IV estimates, similarly to other cases discussed in Rosenzweig and Wolpin (2000).

2. Redshirting in the ECLS-K Data Set

I am not aware of any historical data on redshirting. Thus I cannot draw any conclusions from the above discussion about the validity of the estimates presented in papers such as Angrist and Krueger (1991), Acemoglu and Angrist (2000), or Lochner and Moretti (2004). Nevertheless, a first glance at this issue can be seen in Table 1, which shows the average age at enrollment for boys in the 1960 Census by quarter of birth (as reported in Angrist and Krueger (1991)), including the difference between the census average and the age predicted by our simple assumptions. Note that there is a large discrepancy in Q_4. This difference could be generated by schools with cutoff dates before January 1st, but it could also be evidence of significant redshirting in the fourth quarter.

In the ECLS-K data set we observe four groups of children whose birth dates are within one month of their school's cutoff date, as shown in Figure 1:

Cohort C_1 : Children eligible in the fall of 1997 but who chose to wait until the fall of 1998 to enroll.

Cohort C_2 : Children ineligible in the fall of 1997 who waited until the fall 1998 to enroll.

Cohort C_3 : Children eligible in the fall of 1998 who chose to enroll.

Cohort C_4 : Children ineligible in the fall of 1998 but who chose to enroll anyway.

As shown in Table 3, very few parents would choose to delay their child when first eligible if they are in Cohort C_2. In contrast, cohort C_1 is made up entirely of children whose parents chose to delay their enrollment when eligible in 1997, while C_3 is made up entirely of children whose parents chose to enroll them when eligible in 1998. This allows us to exploit the difference in the composition of groups C_1 and C_3 to obtain information about those children whose parents would choose to delay their enrollment. Tables 4–9 show how these cohorts differ. The children who delayed enrollment were disproportionately wealthy, white, male, English-speaking, and had better-educated parents. These characteristics are correlated with educational attainment, casting doubt on the validity of the instrument generated by entrance cutoff dates and compulsory education laws.

References

Acemoglu, D. and J. Angrist (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual 15*, 9–59.

Aliprantis, D. (2007). When should children start school? *Working Paper*.

Angrist, J. D. and G. W. Imbens (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association 90*(430), 431–442.

Angrist, J. D. and A. B. Krueger (1991). Does compulsory attendance affect schooling and earnings? *The Quarterly Journal of Economics 106*(4), 979–1014.

Bound, J. and D. A. Jaeger (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In S. W. Polachek (Ed.), *Worker Well Being*, Volume 19, pp. 83–108. Research in Labor Economics.

Dobkin, C. and F. Ferreira (2007, July). Do school entry laws affect educational attainment and labor market outcomes? *Working Paper*.

Lochner, L. and E. Moretti (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *The American Economic Review 94*(1), 155–189.

Rosenzweig, M. R. and K. I. Wolpin (2000). Natural "Natural Experiments" in Economics. *Journal of Economic Literature 38*, 827–874.

Tables

Table 1: Entrance Age (in Years) by Quarter of Birth

Entrance Age (in Years)				
Entry	Q_1	Q_2	Q_3	Q_4
When Eligible	6.5	6.25	6	5.75
Redshirted	7.5	7.25	7	6.75
1960 Census Avg	6.45	6.28	6.08	6.07
'60 – Eligible	-0.05	0.03	-0.08	-0.32

Table 2: Years of Schooling Completed before Eligible to Drop Out by Quarter of Birth

Years of Schooling Completed				
Entry	Q_1	Q_2	Q_3	Q_4
When Eligible	9.5	9.75	10	10.25
Redshirted	8.5	8.75	9	9.25

Table 3: Redshirting in the ECLS-K

	Percent Redshirting											
Cohort	$B_1\,(C_1+C_3)$	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9	B_{10}	B_{11}	$B_{12}\,(C_2)$
% Redshirting	28.8	29.8	13.9	14.0	11.3	7.7	8.0	6.8	6.1	5.2	5.4	3.7

Table 4: Descriptive Statistics by Cohort

Different Variables Summarized by Cohort						
Variable	Statistic	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Income	Mean	52,946	61,098	52,456	48,531	61,912
	Std. Dev.	36,375	40,389	34,959	34,466	40,146
	n	12,326	241	702	619	57
# of Books	Mean	74.35	86.89	75.11	69.74	76.84
	Std. Dev.	59.87	60.64	57.57	57.45	65.18
	n	14,910	300	849	744	63
# of Children's Records, Tapes, or CD's	Mean	15.37	18.26	16.09	13.99	17.70
	Std. Dev.	17.86	20.33	18.81	17.42	17.68
	n	15,008	303	851	750	63
Age at entry to K (in months)	Mean	65.56	68.34	70.63	60.07	62.00
	Std. Dev.	4.26	5.53	2.91	1.67	5.25
	n	15,042	302	851	751	63
Mother's Age (At First Birth)	Mean	23.98	25.26	23.93	23.64	24.12
	Std. Dev.	5.44	5.74	5.58	5.37	5.49
	n	14,054	288	793	696	57
Cost Nonparental Childcare (Pre-K, in $)	Mean	43.82	51.43	41.50	39.74	59.38
	Std. Dev.	51.72	49.18	46.17	44.66	77.77
	n	11,465	223	653	559	54
Age at first Nonparental Care (in Months)	Mean	22.28	24.08	22.80	20.69	17.71
	Std. Dev.	19.59	20.95	20.72	18.24	17.71
	n	13,047	266	754	637	58
Hours Spent in Nonparental Care (Pre-K)	Mean	25.23	22.05	26.66	25.33	27.08
	Std. Dev.	21.73	19.79	23.31	22.15	20.32
	n	14,858	298	839	744	63

Table 5: WIC Benefits

Did the Child Receive any WIC Benefits as a Child (in %)					
	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Yes	42.88	31.35	40.66	47.87	39.68
No	56.00	67.66	58.05	51.06	60.32
n	15,064	303	851	752	63

Table 6: Child's Language to Mother

Language to Mother in %					
Speaks Non-English Language	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Never	79.53	93.07	80.85	79.26	77.78
Sometimes, Often, or Very Often	18.53	5.28	16.93	18.36	17.46
n	15,064	303	851	752	63

Table 7: Race

The Composition of Cohorts by Race (in %)					
Race	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
White, Non-Hispanic	56.31	75.29	60.95	53.94	63.01
Black or African-American, Non-Hispanic	14.20	9.30	13.18	13.87	12.33
Hispanic, Race Specified	8.52	3.49	7.51	6.58	1.37
Hispanic, Race Not Specified	8.91	5.52	7.71	11.40	4.11
Asian	6.35	2.03	5.68	7.17	15.07
Native Hawaiian or Other Pacific Islander	1.14	–	1.12	2.47	–
Native American or Native Alaskan	1.80	1.74	1.01	2.00	1.37
More than one race, Non-Hispanic	2.55	2.62	2.64	2.59	2.74
n	17,527	344	984	851	73

Table 8: Gender

The Composition of Cohorts by Gender (in %)					
Gender	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Male	51.15	63.66	53.25	49.82	32.88
Female	48.78	36.34	46.75	50.18	67.12
n	17,554	344	986	851	73

Table 9: Parent's Highest Education Level

Parent's Highest Education Level by Cohort in %					
	ECLS-K	Cohort 1	Cohort 2	Cohort 3	Cohort 4
8th Grade or Below	3.09	0.95	2.76	3.05	1.43
9th-12th Grade	6.55	3.15	6.62	6.62	4.29
High School Diploma/Equivalent	23.80	17.67	25.69	27.61	21.43
Voc/Tech Program	6.77	5.36	8.27	7.12	4.29
Some College	26.50	27.13	25.69	27.86	28.57
Bachelor's Degree	18.53	26.81	16.43	17.05	21.43
Grad/Professional School - No Degree	2.56	3.15	2.65	1.65	2.86
Master's Degree (MA, MS)	7.50	7.89	6.73	6.49	12.86
Doctorate or Professional Degree	4.71	7.89	5.18	2.54	2.86
n	15,961	317	907	786	70

Figures

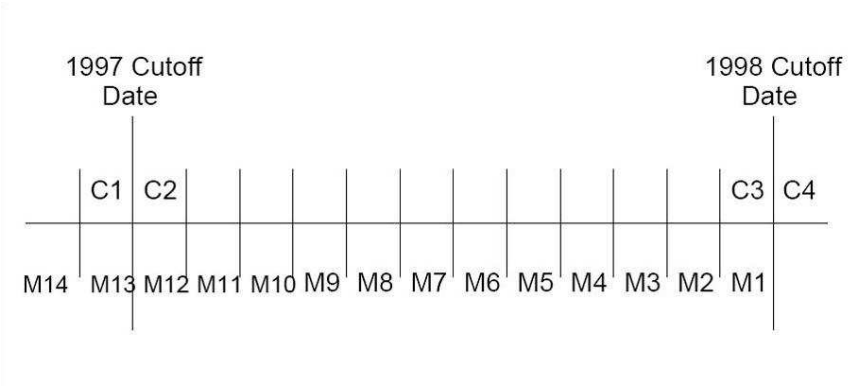

Figure 1: ECLS-K Cohorts