NBER WORKING PAPER SERIES

ON THE VALIDITY OF SEASON OF BIRTH AS AN INSTRUMENT IN WAGE EQUATIONS: A COMMENT ON ANGRIST AND KRUEGER'S "DOES COMPULSORY SCHOOL ATTENDANCE AFFECT SCHOOLING AND EARNINGS?"

John Bound
David A. Jaeger

Working Paper 5835

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1996

The views expressed in this paper are solely the authors' and do not necessarily reflect those of the U.S. Bureau of Labor Statistics or the National Bureau of Economic Research. We thank Robert Margo and Richard Steckel for helpful discussions, and Charlie Brown, Claudia Goldin, Larry Katz, Gary Solon, and participants at the NBER Summer Institute for comments on earlier drafts. This paper is part of NBER's research program in Labor Studies.

On the Validity of Season of Birth as an Instrument in Wage Equations: A Comment on Angrist and Krueger's "Does Compulsory School Attendance Affect Schooling and Earnings?"

John Bound and David A. Jaeger

NBER Working Paper No. 5835
November 1996
JEL Nos. J24
Labor Studies

ABSTRACT

In an important and provocative paper, "Does Compulsory School Attendance Affect Schooling and Earnings?", Angrist and Krueger use quarter of birth as an instrument for educational attainment in wage equations. To support a causal interpretation of their estimates, they argue that compulsory school attendance laws *alone* account for the association between quarter of birth and earnings. In this note we present evidence that the association between quarter of birth and earnings is too strong to be fully explained by compulsory school attendance laws in the samples studied by Angrist and Krueger. Moreover, while the association between quarter of birth and educational attainment was weaker for more recent cohorts, we found no evidence that the strength of the relationship between quarter of birth and earnings was also weaker in those cohorts. In addition, we present evidence that suggests the association between quarter of birth and earnings or other labor market outcomes existed for cohorts that were not bound by compulsory school attendance laws. Our results call into question the validity of any causal inferences based on Angrist and Krueger's estimates regarding the effect of education on earnings.

John Bound
Department of Economics
University of Michigan
Ann Arbor, MI 48109
and NBER
jbound@umich.edu

David A. Jaeger
Bureau of Labor Statistics
Suite 4945
2 Massachusetts Avenue, NE
Washington, DC 20212
jaeger_d@bls.gov

The association between educational attainment and earnings is strong and persistent. Economists and other social scientists have been reluctant, however, to interpret as purely causal the relationship between education and earnings. Educational attainment is not randomly distributed across the population. Rather, individuals who do well in school or are from privileged backgrounds are much more likely to continue their education. While it is possible to write down economic models in which the association between education and earnings understates the causal impact of education on earnings [Griliches 1977, Willis 1986], most economists have presumed that the observed association actually exaggerates the causal effect of educational attainment on earnings.

Even though the endogeneity of educational attainment has long been recognized, plausible instruments for education that would allow for estimation of the causal effect of education on earnings have been difficult to find. In a recent provocative and influential paper, "Does Compulsory School Attendance Affect Schooling and Earnings?" Angrist and Krueger [1991, henceforth AK] present results in which they use information on quarter of birth to form instruments for educational attainment. AK's instrumental variables (IV) estimates tend to be close to and often larger than their ordinary least squares (OLS) estimates. They interpret these results as "cast[ing] doubt on the importance of omitted variables bias in OLS estimates of the return to education ..."

In this paper we review evidence that we believe calls into question the validity of any causal inferences based on estimates in which information on quarter of birth is used to construct instruments for educational attainment in earnings equations. We first review AK's framework. In Section 2 we present evidence suggesting that the associations between quarter of birth and both educational attainment and earnings are too strong to be explained solely by the compulsory school attendance law mechanisms detailed by AK. In Section 3 we present results that suggest that an association between quarter of birth and labor market outcomes predated the

enactment of compulsory school attendance laws in the U.S. In Section 4, we review the large literature documenting associations between quarter of birth and a host of factors that could plausibly affect earnings. Finally, in Section 5 we review the validity of arguments that Angrist and Krueger use to support the use of quarter of birth to form instruments for educational attainment.

1. Angrist and Krueger's Framework

AK find significant associations between quarter of birth, educational attainment, and earnings for cohorts of men born during the 1920s, 1930s and 1940s. In their data, men born during the first quarter of the year obtain less education and have lower earnings than those born during the remainder of the year. AK argue that these associations are due to compulsory school attendance laws. The typical law requires students to start first grade in the autumn of the calendar year in which they turn 6 and to stay in school until their 16th birthday. Thus, individuals born in the early months of the year will usually enter first grade when they are near seven years of age and will turn 16 in the middle of tenth grade. Individuals born in the third or fourth quarter will typically start school either just before or just after their sixth birthday and will finish tenth grade before they turn 16.

AK present several tabulations in support of their assertion that compulsory school attendance laws are partially responsible for the relationship between quarter of birth and educational attainment. First, the observed relationship between educational attainment and season of birth is weaker in more recent cohorts that were less likely to have been constrained by the law. Second, they find a weaker relationship between quarter of birth and education for individuals with higher levels of education. Third, among individuals who recently turned 16, enrollment rates are higher in states that require students to remain in school through their 17th or 18th birthday than in states that allow 16 year olds to leave. Each of these find-

ings is consistent with the argument that compulsory school attendance laws are responsible for an association between quarter of birth and educational attainment.

AK leave us with little doubt that compulsory school attendance laws are at least partially responsible for the observed correlation between quarter of birth and educational attainment. We are skeptical, however, that these laws are the *only* reason for this correlation, and, therefore, whether quarter of birth is a legitimate instrument in estimating wage equations.

2. The Association between Quarter of Birth, Education, and Earnings

Quarter of Birth and Completion of the Highest Grade Attended

Compulsory school attendance laws typically require students born during the summer to finish tenth grade, while permitting those born during the winter to drop out in the middle of the tenth grade. If compulsory schooling laws were the only reason for the relationship between quarter of birth and educational attainment, we would expect that most of the association would be accounted for by a larger fraction of individuals born in the summer completing the highest grade they attended. Moreover, we would predict an association between quarter of birth and educational attainment only for individuals with a high school education or less, but would not expect any association between season of birth and educational attainment above the high school level.

In Table 1, we use the same samples as AK[1] and report the fraction of men born during the 1920s, 1930s and 1940s that did and did not complete the highest grade of school they attended. The residual category includes both respondents who reported that they were attending school at the time of the Census as well

[1] Using their Appendix 1, we were able to exactly replicate AK's 1930-49 and 1940-49 birth cohort samples from the 1980 Census. We were unable to exactly replicate their 1920-29 birth cohort from the 1970 Census, however, using information from the same Appendix. Our sample is slightly smaller than theirs (245,299 versus 247,199).

as a negligible portion that reported never having attended school. The χ^2 tests reported are for the full 4×4 table of quarter of birth $\times$ each of the four responses to the grade completion question. The association between grade completion rates and quarter of birth is very weak. For the cohorts born during the 1920s and 1930s, individuals born during the winter months were slightly less likely to finish the highest grade of school they attended, but the magnitude of the difference is trivial and accounts for little of the difference in the total number of years of education completed. For the cohorts born during the 1940s, men born in the winter quarter were actually more likely to have finished the highest grade they attended.

Men born during the 1940s would have been in high school during the late 1950s and early 1960s. It is possible, therefore, to use the 1960 Census to directly examine the effect of compulsory school attendance laws on the high school enrollment patterns for this cohort. Restricting attention to men born in 1944 (who would turn 16 some time during 1960) and residing in states that required students to attend school through their 16th birthday, we calculate that the fraction of those that had already turned 16 who were still enrolled as of the April 1, 1960 Census date was 0.894. For those that had yet to turn 16, the fraction was higher: 0.930. Together, these numbers suggest that, of those that turned 16 during the school year, roughly 3.5 percent (standard error of .7) dropped out without finishing the grade they were in. This 3.5 percent is too small to account for the approximately 0.07 years difference in educational attainment between those born during the first and the third quarter for men born during the 1940s.[2]

In contrast to what we would expect if compulsory school attendance laws fully accounted for the association between quarter of birth and educational attainment, there also appears to be an association between quarter of birth and post-secondary

[2] AK report similar tabulations in Table II of their paper. For men and women combined they estimate that for the cohort born during 1944, compuslory school attendance laws were binding on about 4 percent of the population.

educational attainment. AK report that those born in the first quarter of the year are between .4 and .7 percent less likely to finish college than those born later in the year, while Angrist and Imbens [1995] show that the entire distribution of educational attainment is shifted downward for those born in the first quarter.

If compulsory school attendance laws not only induce individuals born in the summer to complete 10th grade, but also induce them to continue their education past 10th grade, such laws could still fully account for the observed differences in educational attainment across seasons of birth. While we do not find such effects to be very plausible, we can not rule them out as impossible. It is worth noting, however, that the interpretation of AK's results is affected if compulsory school attendance laws did not simply induce men to stay in school until they were 16. Under the assumption that information on quarter of birth is a valid instrument (i.e. under the assumption that there is no direct association between quarter of birth and earnings), the estimates AK report are a weighted average of the effect of education on earnings for those induced by the compulsory attendance laws to continue their education, with the weights being proportional to years of education gained [Angrist and Imbens 1995]. To illustrate the implication for AK's results, suppose that compulsory attendance laws raise the educational attainment for some individuals born late in the year by negligible amounts, but raises the educational attainment for others born at the same time by years. If those induced to obtain several additional years of education tended to be those with higher returns, then the estimated average effect would overestimate the actual average value of an extra year of education for those affected by the compulsory attendance law.

Quarter of Birth and Earnings

As we noted earlier, AK find that the association between quarter of birth and educational attainment is substantially weaker for cohorts born more recently. If compulsory school attendance laws are the only cause for the association between

quarter of birth and earnings, we would expect a similar pattern for earnings. The opposite appears to be the case, however.

Table 2 reports estimates of the effect of quarter of birth on educational attainment and the logarithm of weekly earnings for AK's samples.[3] The sum of the quarter of birth coefficients has been constrained to sum to zero. Because there is an association between age and educational attainment and earnings and because, in a cross section, quarter of birth and age are related by definition, it is important to carefully control for age when estimating the effect of quarter of birth on educational attainment on earnings. It seems natural to assume that earnings are a continuous and smooth function of age and therefore to model this effect using polynomials of age.[4] On the other hand, the effects of compulsory school attendance laws will induce an association between age and earnings that includes discrete jumps in earnings between those born in the fourth quarter of one year and those born in the first quarter of the next. It is, in fact, differences in the pattern of expected impacts that allow for separate identification of the effect of age and quarter of birth on education and earnings. In addition to the quarter of birth estimates themselves, we report two measures of the overall strength of the association between quarter of birth and education and earnings: the difference between the coefficients on the third and first quarters and the sum of the absolute values of the coefficients on all four quarter of birth indicators.

Consistent with the results reported by AK, Table 2 shows that the strength of the association between quarter of birth and educational attainment declines considerably between cohorts born in the 1920s and those born in the 1940s. In

[3] We employ the same control variables as AK use in much of their analysis: indicators for black, married with spouse present, lives in an SMSA, and 8 regions of residence.

[4] In the results we report, we use a quadratic, but have experimented with cubics and quartics. In no case did the inclusion of these higher order terms have any significant effect on the estimated association between quarter of birth and either educational attainment or earnings. These results are available from the authors.

contrast, the strength of the association between quarter of birth and earnings is actually somewhat larger for more recently-born cohorts. Figure 1 shows the third quarter-first quarter difference for both educational attainment and log weekly earnings, and makes clear the change in the strength of these relationships across AK's three birth cohorts. The increase in the strength of the association between quarter of birth and earnings suggests that factors other than compulsory school attendance laws must explain the association between quarter of birth and earnings. Moreover, the fact that the strength of the association between quarter of birth and earnings increased while the strength of the association between quarter of birth and educational attainment declined is indicative of a possible direct association between quarter of birth and earnings.[5]

Instrumental Variables Estimates of the Returns to Education

The rise in the strength of the association between education and earnings relative to that between quarter of birth and earnings has clear implications for instrumental variable estimates of the returns to education. Table 3 explores this issue using AK's three samples. Columns (1), (3) and (6) contain OLS estimates while the remaining columns present results from a variety of IV specifications that include both different age controls and different instruments for education.[6] As indicated, columns (3), (5), (6), and (8) replicate estimates reported by AK in columns (5) through (8), Tables IV through VI. In addition to the coefficients on educational attainment and their standard errors, we also report the partial R^2 of the excluded instruments and F statistic from the test of statistical significance of the excluded instruments, both from the first stage regression, for each IV specification.

[5] Dramatic rises in the returns to education between 1970 and 1980 could explain the rising quarter of birth differentials observed in Table 2. We have good reason to believe, however, that the returns to education actually fell over this period [Freeman 1975, Katz and Murphy 1992].

[6] Other control variables are the same as in Table 2.

The simplest and perhaps most plausible specification is shown in column (2). This model controls for age using a simple quadratic and uses three quarter of birth dummies as the excluded instruments. While the OLS estimates for any of the specifications remain quite stable across cohorts, the IV estimates vary substantially, more than tripling between the cohort born in the 1920s and that born in the 1940s. This dramatic increase in these estimates can be accounted for by the substantial weakening of the relationship between quarter of birth and educational attainment while the strength of the relationship between earnings and quarter of birth increased. The IV estimates from this model for cohorts born during the 1930s and 1940s are implausibly large.[7]

Column (7) adds year of birth dummies to the specification of column (2), yielding large effects on both the coefficient and standard error estimates. This implies that much of the identification for the coefficients in column (2) derives from variation in the effects of quarter of birth across single-year cohorts. Adding quarter of birth $\times$ year of birth interactions as instruments (column (8)) increases the precision of the estimates substantially. While there is some theoretical justification for including these interaction terms, we are concerned that most of the identifying information on the effect of education comes from them. With quarter of birth $\times$ year of birth interactions included, the reported F statistics on the excluded instruments in the first stage are small enough to suggest that quantitatively important finite sample biases may affect the estimates (see Bound, Jaeger, and Baker [1995] as well as Staiger and Stock [forthcoming]). In addition, the standard errors on the specifications that include a quadratic in age as control variables are large enough to imply that the IV estimates have very little power to distinguish between plausible alternative point estimates.

[7] Mayer and Knutson [1996] find similar implausibly high rates of return for more recent cohorts.

AK do not report our column (2) estimates, but control for the effect of age on earnings by either employing 9 year of birth indicators or including these indicators as well as a quadratic in age. The specifications that include only year of birth dummies are problematic, as they do not control for age variation within birth year, despite the fact that age (measured in quarter years) and quarter of birth are perfectly collinear within single-year birth cohorts. While this specification does not seem very sensible to us, for comparison purposes we present results from this model in columns (4) and (5).

3. Quarter of Birth and Labor Market Outcomes for Men Born 1840-1875

If compulsory school attendance laws do, in fact, account for all of the association between quarter of birth and both education and earnings, we would expect not to observe such associations in cohorts educated before such laws were enacted.[8] Most states enacted their first compulsory attendance laws between 1870 and 1915 [Eisenberg 1988]. Initial laws usually required school attendance through at least age 14.

To test the hypothesis that no relationship exists between quarter of birth and earnings or other labor market outcomes in cohorts educated before the enactment of compulsory school attendance laws, we employ data from the 1900 Census. The public use files from this Census provide both the month and year of birth. While data on educational attainment and earnings are not available in the 1900 Census, information on the nature of work performed by the individuals in the sample has been recoded to 1950 occupational codings. To crudely proxy 1900 earnings, we calculated the mean log weekly earnings by occupation in the 1950 Census and used those means to impute earnings data to the 1900 microdata.[9] These mean

[8] This idea arose in a conversation between Peter Klenow and Gary Solon. We are indebted to Gary Solon for mentioning it to us.

[9] Only "sample line" individuals had information on earnings in the 1950 Census. Individuals whose sex, age, race, earnings, or occupation were allocated or who had unknown or missing

earnings were then matched to the microdata from 1900 by 1950 occupation code. The samples from both Censuses are limited to native-born white men.

In columns (1) through (4) of Table 4 we present reduced form regressions of (imputed) log weekly earnings on four quarter of birth dummy variables (with coefficients constrained to sum to zero), and a quadratic in age (measured in months). Columns (2) and (4) also include demographic control variables similar to those used above: indicator variables for eight regions of residence, married with spouse present, and living in a city with more than 50,000 in population.[10] The first two columns are estimated using a sample of white native men born in 1840-55 who were neither living nor born in Massachusetts or Vermont, the two states that enacted compulsory schooling laws before 1870. Virtually no one represented by this restricted sample would have been subject to compulsory school attendance laws while growing up. The resulting sample is quite small, however. In the third and fourth columns we add to the original sample white native men born in 1856-1875 who would not have been affected by compulsory schooling laws.[11] Because the imputation procedure induces heteroskedasticity in the error terms of the models,

---

occupational data were dropped from the 1950 sample. Categorical earnings data are given values in the midpoint of the interval of the category.

[10] Only those individuals with a valid month of birth and valid occupational code were kept in the samples. Individuals living in "Indian territories" or military reservations were dropped from the samples.

[11] Men were excluded from this sample if they might have been subject to compulsory school attendance laws, according to their year of birth, their state of birth, and their state of residence in 1900. For example, Ohio enacted a law in 1877 requiring individuals to attend school through their fourteenth birthday. We therefore excluded from the sample men born after 1862 if they were born in Ohio or lived there in 1900. This algorithm eliminates many, but not all, of those men bound by compulsory school attendance laws. In particular, our algorithm would not eliminate men who grew up in states that enacted compulsory school attendance laws early but who were born and eventually settled in states that were late to adopt such laws. During the 19th century migration was primarily east to west [Steckel 1983, 1989], however, and states in the Midwest tended to adopt compulsory school attendance laws later than those in the Northeast [Eisenberg 1988]. It seems unlikely, therefore, that there are many men constrained by compulsory schooling laws who were not excluded by our algorithm.

we present heteroskedasticity-consistent standard errors estimated using White's (1980) method.

Despite the fact that the men represented by these two samples from the 1900 Census were unlikely to be affected by compulsory school attendance laws, their imputed earnings show a distinct pattern of seasonality, with wages being higher for those born in the summer than for those born in the winter.[12] This pattern is similar to the pattern for men born during the first half of the 20th century, but, as the summary measures indicate, the amplitude of these effects were substantially larger for the earlier cohorts.

More than half of these samples were employed in agriculture and the relationship between quarter of birth and the probability of being an agricultural worker drives, in part, the results in columns (1) through (4). In columns (5) through (8) of Table 4 we present logit estimates of the association between quarter of birth and employment in agriculture for these two samples. These estimates show that a winter birth raised the odds that a man was employed in agriculture by more than 10 percent.

We are cautious about overinterpreting these results from the 1900 Census, for several reasons. First, the factors at work in creating an association between labor market outcomes during the 19th century might not be operative or as important during the 20th century.[13] Second, the labor market 100 years ago was very different than it was in 1970 and 1980. It is possible that the factors inducing an association between quarter of birth and labor market outcomes might have had

[12] The large range in age in these samples (particularly the 1840-1875 cohort) made us particularly concerned about how we controlled for age in these models. In addition to parametrically controlling for age with a quartic, we also estimated generalized additive models [Hastie and Tibshirani 1990] in which age was modeled as a locally-weighted running line smooth (loess) and the other variables in the model were entered parametrically. We set the span of the smooth so that it had approximate degrees of freedom of 5. Both the quartic and loess models produced estimates of the quarter of birth coefficients that were virtually identical to those presented in Table 4.

[13] For example, the decline in infant mortality would have lessened the potential for selective survivorship to explain seasonal patterns in labor market outcomes.

a more important role in the labor market in 1900 than in the latter part of the century. Nevertheless, unless the factors at work at the the turn of the century were completely inoperative 70 and 80 years later, our results imply that the association between quarter of birth and labor market outcomes is not simply due to the effect of compulsory schooling attendance laws.

4. The Association between Quarter of Birth and Various Factors

We have argued that the available evidence suggests compulsory school attendance laws are not the only factor behind the association between quarter of birth and either educational attainment or earnings. Nothing in these results proves, however, that there is a direct association between quarter of birth and earnings. While we know of no incontrovertible evidence on the direct effect of quarter of birth on earnings, it seems quite conceivable that such effects exist. Researchers have documented associations between quarter of birth and a variety of factors that either are known to affect earnings or might plausibly do so. We briefly outline some of these factors below.

Performance in school. There is some evidence that quarter of birth is related to performance in school. School attendance rates [Carroll 1992], the likelihood that a student will be evaluated as having behavioral difficulties [Mortimore *et al.* 1988], the probability that a student will be referred for mental health services [Tarnowski, *et al* 1990], and performance in reading, writing, and arithmetic [Mortimore *et al.* 1988; Williams *et al.* 1970 summarizes the earlier literature] have all been shown to vary by quarter of birth. The evidence regarding differences in IQ of children born at different times of the year [Whorton and Karnes 1981] is somewhat inconclusive.

Health differences. There is a substantial literature documenting that individuals born early in the year are more likely to suffer from schizophrenia (see, for example, Watson *et al.* [1984], O'Callaghan *et al.* [1991], and Sham *et al.* [1992]). Other health factors that vary by quarter of birth are the incidence of mental retardation

[Knoblock and Pasamanick 1958], autism [Gillberg 1990], dyslexia [Livingston *et al.* 1993], multiple sclerosis [Templer *et al.* 1991], and manic depression [Hare 1975].

Regional patterns. Clear regional patterns in birth seasonality show that there is substantially more seasonal variation in births in the southern states [Lam and Miron 1991]. For example, using data on white births in Georgia between 1948 and 1986, they found that there were 16 percent more births in September than in May. In contrast, whites were only 4 percent more likely to be born in September than in May in New York during the same period.

Race. Lam and Miron [1991] also document clear racial patterns in birth seasonality in the U.S., with blacks showing more seasonal variation in births than whites. For example, they find that (nationally) whites are 11 percent more likely to be born in September than May, compared to 16 percent for nonwhites.[14] These patterns cannot be explained by the overrepresentation of blacks in the South, as racial differences persist when Lam and Miron focus on specific states.

Family Income. Kestenbaum [1987], using the 1980 Census, reports simple bivariate associations showing that children born to families with high incomes are more likely to be born in the spring. In Table 5 we confirm Kestenbaum's finding in a multivariate setting using data drawn from several Censuses and birth cohorts.[15] The table presents results of regressions of log per capita family income on quarter of birth dummy variables, with the sum of the coefficients constrained to be equal to zero. The second, fourth, and sixth columns add controls for race and geographic region of residence to the models. In general, we find that children born in the

[14] These results are for first births between 1960 and 1986. A similar pattern holds for second births. Lam and Miron do not report results for all or higher order births for whites and nonwhites in the U.S.

[15] Per capita family income is defined as the total household income divided by the number of persons in the household. Income is a categorical variable in the 1960 and 1970 public use files. In those data we use the midpoint of the interval represented by the category.

second quarter tend to be born into families with incomes that are statistically significantly above average, even after controlling for race and region of residence.

Personality. Personality traits also may vary by season of birth. Using data from the National Longitudinal Survey of Youth, Gortmaker *et al.* [forthcoming] report a strong association between shyness and month of birth, with those born in the spring being 50 percent more likely to suffer extreme shyness.

It cannot reasonably be disputed that there are relatively strong and robust associations between quarter of birth and various social outcomes. The degree of bias that these associations induce in the estimated return to education is less clear. The very weak relationship between quarter of birth and education attainment indicates, however, that even minor associations between quarter of birth and other seasonal factors could produce large effects on the estimated coefficients [Bound, Jaeger, and Baker 1995]. Some phenomena, such as schizophrenia, affect a very small number of individuals in the population and would not be expected to account for much of the observed association between season of birth and education and earnings. Other factors, such as race, geographic region of residence, and family background are potentially much more important.

To illustrate, using a specification that includes age and age squared as control variables while instrumenting for educational attainment with three quarter of birth indicators, we find that the addition of the "black" dummy variable lowers the estimated return to education by between 0.6 and 1.7 percentage points in the three samples from AK. Using the same base specification, inclusion of 8 region of birth indicators affects the estimated return by between -0.1 and 2.2 percentage points. It is also possible to crudely estimate the impact of leaving out family income from the IV estimates. Using the seasonal patterns of per capita family income reported in Table 4 in conjunction with this base specification estimated in AK's three birth cohorts, we crudely estimate the inclusion of family income would actually raise the estimated return from anywhere between 0.1 and 6.6 percentage points, depending

on which birth cohort and Table 4 results were used.[16] These calculations certainly suggest that the seasonal patterns in some of the factors discussed above are sufficiently large to have a quantitatively important, if unpredictable, impact on IV estimates of the return to education in which quarter of birth is employed as an instrument for educational attainment.

5. Angrist and Krueger's Arguments for the Exogeneity of Quarter of Birth

AK are aware that the validity of their estimates of the causal effect of education on earnings rests on the assumption that there is no direct association between quarter of birth and earnings, and they present a number of arguments to support this notion in Section III of their paper.

In their discussion AK specifically discuss a number of the factors mentioned in Section 4 of this paper. In particular, they briefly discuss the effect of age at school entry on educational achievement, noting that research in the area is weak and inconclusive. We feel that this characterization of the literature is accurate, but we do not think this fact validates the use of information on quarter of birth to form instruments for educational attainment. Children mature a good deal between the age of 5 3/4 and 6 3/4. Thus, it seems plausible that the exact age at school entry (and thus quarter of birth) could affect school performance and potentially earnings. Weak and inconclusive research evidence on the subject does not imply that quarter of birth does not affect school performance or earnings, but rather that the direction and magnitude of such effects are unknown. This is a comfort only if we take the position that instruments are valid unless proven otherwise.

[16] In this calculation, we assumed that a 1 percent rise in family income would raise a son's weekly earnings by 0.4 percent and his educational attainment by 0.014 years. The 0.4 percent figure is based on Solon's [1992] finding that the intergenerational correlation in income between fathers and sons was about 0.4. The 0.014 year figure represents an unreported result based on the same samples used in Solon [1992].

AK also discuss possible correlations between quarter of birth and parental socioeconomic status. In discussing this issue AK summarize the (at the time) unpublished work of Lam and Miron [1987]:

> ...[They] present a variety of evidence suggesting that season of birth is unrelated to the socioeconomic status (and other characteristics) of the parents... [T]hey find that the seasonal pattern of birth is similar ... across regions of the United States ..., and within countries before and after dramatic economic transitions (p. 1008).

This summary of Lam and Miron's work is at odds with their results that we reported in Section 4. Lam and Miron do find a summer peak and a winter trough in births for various U.S. populations, but (as we noted above) the amplitude of these cycles varies considerably across the population.

AK argue that explanations for the association between quarter of birth and educational attainment and earnings that rely on such things as age at school entry or parental socioeconomic factors cannot explain the decline over time in the association between quarter of birth and educational attainment. Depending on the mechanisms involved, it is not hard to imagine explanations for why the association between season of birth and educational attainment might have declined, even if it were due to the effects of compulsory schooling attendance laws. Moreover, birth seasonality has declined dramatically over time in a number of countries, including the U.S. [Lam and Miron 1991, Seiver 1985]. Thus, it would seem plausible that the effect of season of birth on various outcomes would also have changed over time as well. Regardless, the evidence on the declining association between quarter of birth and educational attainment *does not* imply that compulsory school attendance laws are the *only* reason for the association between quarter of birth and educational attainment.

Towards the end of their Section III, AK present what would appear to be simple and clear evidence that there is no direct association between quarter of birth and earnings. Some of this evidence rests on problematic assumptions, however.

AK report that, in their samples, the coefficients on quarter of birth indicators in regressions of earnings on education are statistically insignificant. [17] But this is a valid test of the exclusion restrictions necessary for quarter of birth indicators to be valid instruments in wage equations only if education itself is exogenous! Similarly, they report that while there is an association between quarter of birth and earnings in samples of men at all levels of educational attainment, no such association exists for college graduates. As long as education is endogenous, however, stratifying on education is problematic, introducing selection bias in the estimates from the self-selected sample.

6. Conclusion

We have presented evidence that calls into question the notion that compulsory school attendance laws are the only reason for the observed association between quarter of birth and educational attainment and earnings. For cohorts of men born during the first half of the 20th century, the association between quarter of birth and both educational attainment and earnings would appear to be too strong to be explained solely by compulsory school attendance laws. Moreover, we find an association between quarter of birth and labor market outcomes in cohorts whose education predates the existence of effective compulsory attendance laws.

Even if compulsory schooling attendance laws cannot account for the strong associations between quarter of birth and either educational attainment or earnings, this does not necessarily imply a direct association between quarter of birth and earnings. However, we have summarized a large literature documenting strong associations between quarter of birth and a variety of factors that could plausibly have direct effects on earnings. Moreover, IV estimates of the effect of education on

[17] In contrast, we find that when quarter of birth indicator variables are added to the OLS specifications in Table 3 they often have statistically significant coefficients. For example, a test of the joint significance of two quarter of birth indicator variables added to the specification in Table 3, column 6 for the 1940-49 cohort yields and F statistic of 5.965.

earnings appear to be sensitive to the details of the specification used. Estimates based on what is, in our opinion, the most defensible model specification give implausibly large estimates of the return to education in two of the three birth cohorts that AK examined.

Perhaps the most striking result we presented is that the association between quarter of birth and educational attainment has declined (between cohorts born in the 1920s and those born during the 1940s) while at the same time no similar decline occurred for the association between quarter of birth and earnings. This finding alone is inconsistent with the notion that compulsory schooling attendance laws account for the association between quarter of birth and earnings would appear to strongly suggest a direct association between season of birth and earnings, independent of the effect through education.

Our results do not conclusively *prove* that any possible direct association between quarter of birth and earnings is strong enough to seriously bias AK's IV estimates. Nevertheless, we believe that the evidence presented here, taken as a whole, makes it difficult to have any confidence in the validity of causal inferences drawn from the estimation of wage equations in which information on quarter of birth is used to instrument for educational attainment.

REFERENCES

Angrist, J. D., and G. Imbens (1995) "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, **90**, 431–42.

Angrist, J. D. and A. B. Krueger (1991) " Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, **106**, 979–1014.

Bound, J., D. A. Jaeger, and R. M. Baker (1995) "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, **90**, 443–450.

Carroll, H. C. M. (1992) "Season of Birth and School Attendance," *British Journal of Educational Psychology*, **62**, 391–396.

Eisenberg, M. J. (1988) "Compulsory Attendance Legislation in America, 1870 to 1915," Ph.D. Dissertation, University of Pennsylvania.

Freeman, R. B. (1975) "Overinvestment in College Training," *Journal of Human Resources*, **10**, 287–311.

Gillberg, C. (1990) "Do Children with Autism Have March Birthdays?," *Acta Psychiatrica Scandinavica*, **82**, 152–156.

Gortmaker, S. L., J. Kagan, A. Caspi, and P. A. Silva (forthcoming) "Daylength During Pregnancy and Shyness in Children: Results from Northern and Southern Hemispheres," *Developmental Psychobiology*, , .

Griliches, Z. (1977) "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, **45**, 1–22.

Hare, E. H. (1975) "Manic-Depressive Psychosis and Season of Birth," *Acta Psychiatrica Scandinavica*, **52**, 69–79.

Hastie, T. J., R. J. Tibshirani (1990) *Generalized Additive Models.* London: Chapman and Hall.

Kestenbaum, B. (1987) "Seasonality of Birth: Two Findings from the Decennial Census," *Social Biology*, **34**, 244–248.

Knoblock H., and B. Passamanick (1958) "Seasonal Variation in the Births of the Mentally Deficient," *American Journal of Public Health*, **48**, 1201– 1208.

Lam, D., and J. A. Miron (1987) "The Seasonality of Births in Human Populations," Population Studies Research Report No. 87–114, University of Michigan.

Lam, D., and J. A. Miron (1991) "Seasonality of Births in Human Populations," *Social Biology*, **38**, 51–78.

Livingston, R. B., S. Adam, and H. S. Bracha (1993) "Season of Birth and Neurodevelopmental Disorders: Summer Birth is Associated with Dyslexia," *Journal of the American Academy of Child and Adolescent Psychiatry*, **32**, 612–616.

Mayer, S., and D. Knutson (1996) "Early Education Versus More Education: Does the Timing of Education Matter?," Harris School, University of Chicago, unpublished draft.

Mortimore, P., P. Sammons, L. Stoll, D. Lewis, and R. Ecob (1988) *School Matters: The Junior Years*. Somerset, U.K.: Open Books.

O'Callaghan, E., T. Gibson, H. A. Colohan, D. Walshe, P. Buckley, C. Larkin, and J. L. Waddington (1991) "Season of Birth in Schizophrenia: Evidence for Confinement of an Excess of Winter Births to Patients Without a Family History of Mental Disorder," *British Journal of Psychiatry*, **158**, 764–769.

Seiver, D. A. (1985) "Trend and Variation in the Seasonality of United States Fertility, 1947–1976," *Demography*, **22**, 89–100.

Sham, P. C., E. O'Callaghan, N. Takei, G. K. Murray, E. H. Hare, and R. M. Murray (1992) "Schizophrenia Following Prenatal Exposure to Influenza Epidemics Between 1930 and 1960," *British Journal of Psychiatry*, **160**, 461-466.

Solon, G. (1992) "Intergenerational Income Mobility in the United States," *American Economic Review*, **82**, 393–408.

Staiger, D., and Stock, J. H. (forthcoming) "Instrumental Variables Regression with Weak Instruments," *Econometrica*.

Steckel, Richard H. (1983) "The Economic Foundations of East–West Migration during the 19th Century," *Explorations in Economic History*, **20**, 14–36.

Steckel, Richard H. (1989) "Household Migrations and Rural Settlement in the United States, 1850–1860," *Explorations in Economic History*, **26**, 190–218.

Tarnowski, K. J., D. F. Anderson, R. S. Drabman, and P. A. Kelly (1990) "Disproportionate Referrals for Child Academic/Behavior Problems: Replication and Extension," *Journal of Consulting and Clinical Psychology*, **58**, 240–243.

Templer, D. I., N. H. Trent, D. A. Spencer, A. Trent, M. D. Corgiat, P. B. Mortensen, and M. Gorton (1991) "Season of Birth in Multiple Sclerosis," *Acta Neurologica Scandinavica*, **85**, 107–109.

Watson, C. G., T. Kucala, C. Tilleskjor, and L. Jacobs (1984) "Schizophrenic Birth Seasonality in Relation to the Incidence of Infectious Diseases and Temperature Extremes," *Archives of General Psychiatry*, **41**, 85–90.

White, H. (1980) "Heteroskedasticity-Consistent Covariance Matrix Estimation and a Direct Test for Heteroskedasticity," *Econometrica*, **48**, 817–838.

Whorton, J. E. and F. A. Kárnes (1981) "Season of Birth and Intelligence in Samples of Exceptional Children," *Psychological Reports*, **49**, 649– 650.

Williams, P., P. Davies, R. Evans, and N. Ferguson (1970) "Season of Birth and Cognitive Development," *Nature*, **228**, 1033–1036.

Willis, R. J. (1986) "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics*, ed. Ashenfelter, O., and R. Layard. Amsterdam: North Holland, 525– 602.

Table 1
Completion of Highest Grade Attended
by Quarter of Birth

	Men Born 1920-29 (1970 Census)		Men Born 1930-39 (1980 Census)		Men Born 1940-49 (1980 Census)	
Quarter of Birth	Comp.	Not Comp.	Comp.	Not Comp.	Comp.	Not Comp.
Jan. -- Mar.	.812	.177	.835	.139	.806	.139
Apr. -- June	.814	.174	.837	.137	.805	.138
July -- Sep.	.816	.171	.838	.136	.806	.136
Oct. -- Dec.	.814	.174	.837	.136	.801	.139
χ^2	14.439 [.108]		14.350 [.110]		42.538 [.000]	
N	245,299		329,509		486,926	

NOTE: *p* values in brackets.

Table 2
Reduced Form: Quarter of Birth Effects on Log Weekly Earnings and Educational Attainment

	Educational Attainment			Log Weekly Earnings		
Qtr of Birth	Men Born 1920-29	Men Born 1930-39	Men Born 1940-49	Men Born 1920-29	Men Born 1930-39	Men Born 1940-49
Jan. -- Mar.	-.062	-.049	-.028	-.004	-.007	-.007
	(.011)	(.010)	(.007)	(.002)	(.002)	(.002)
Apr. -- June	-.047	-.015	-.000	-.003	-.004	-.001
	(.011)	(.010)	(.007)	(.002)	(.002)	(.002)
July -- Sep.	.049	.020	.005	.004	.005	.006
	(.011)	(.009)	(.007)	(.002)	(.002)	(.002)
Oct. -- Dec.	.060	.045	.023	.002	.005	.002
	(.011)	(.010)	(.007)	(.002)	(.002)	(.002)
$Q_3 - Q_1$	.111	.069	.033	.008	.012	.013
	(.018)	(.016)	(.012)	(.003)	(.003)	(.003)
$S\|Q_i\|$	.218	.129	.056	.013	.021	.016
	(.024)	(.020)	(.015)	(.004)	(.004)	(.003)
F	23.660	13.486	6.256	2.436	7.154	9.661
	[.000]	[.000]	[.000]	[.063]	[.000]	[.000]
N	245,299	329,509	486,926	245,299	329,509	486,926

SOURCE: 1920-29: 1970 Census, 1930-39 and 1940-49: 1980 Census.
NOTE: All models estimated with OLS. Standard errors in parentheses, *p* values in brackets. *F* is for test of equality of quarter of birth dummy variables. Coefficients on quarter of birth dummy variables are restricted to sum to zero. Age is measured in quarter years. Demographic control variables are lives in central city, married (1=married, living with spouse), and 8 regional dummies.

Table 3
Estimated Return to Education

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) IV	(6) OLS	(7) IV	(8) IV
				Men Born 1920-29 (1970 Census)				
AK Table (Column)			IV(5)		IV(6)	IV(7)		IV(8)
Coefficient	.070	.059	.070	.061	.066	.070	.090	.089
	(.000)	(.021)	(.000)	(.017)	(.015)	(.000)	(.062)	(.033)
F		23.660		37.078	4.415		4.171	1.043
Partial R^2 (x 100)		.029		.045	.054		.003	.012
				Men Born 1930-39 (1980 Census)				
AK Table (Column)			V(5)		V(6)	V(7)		V(8)
Coefficient	.063	.142	.063	.099	.081	.063	.677	.060
	(.000)	(.033)	(.000)	(.021)	(.016)	(.000)	(1.536)	(.029)
F		13.486		30.526	4.747		.088	1.613
Partial R^2 (x 100)		.012		.028	.043		.000	.014
				Men Born 1940-49 (1980 Census)				
AK Table (Column)			VI(5)		VI(6)	VI(7)		VI(8)
Coefficient	.052	.201	.052	-.073	.039	.052	-.116	.078
	(.000)	(.059)	(.000)	(.027)	(.014)	(.000)	(.255)	(.024)
F		6.256		26.315	6.849		.546	2.736
Partial R^2 (x 100)		.004		.016	.042		.000	.016
Age Controls								
Age, Age^2	x	x				x	x	x
9 Yr of Birth Dummies			x	x	x	x	x	x
Instruments								
Qtr of Birth		x		x	x		x	x
Qtr of Birth x Yr of Birth					x			x

NOTE: Standard errors in parentheses. F and partial R^2 are for instruments in first stage. All specifications include Race (1=black), SMSA (1=central city), Married (1=married, living with spouse), and 8 Regional dummies as control variables. Sample sizes are 245,299 for men born 1920-29; 329,509 for men born 1930-39; and 486,926 for men born 1940-49.

Table 4
Reduced Form: Quarter of Birth Effects on
Imputed Log Weekly Earnings and I(Agriculture) for White Men
Educated Prior to Compulsory Schooling Laws

	OLS: Imputed Log Weekly Earnings				Logit: I(Agriculture)			
	Men Born 1840-55 (1900 Census)		Men Born 1840-75 (1900 Census)		Men Born 1840-55 (1900 Census)		Men Born 1840-75 (1900 Census)	
Qtr of Birth	Age only	Age & Demo.	Age only	Age & Demo.	Age only	Age & Demo.	Age only	Age & Demo.
Jan. -- Mar.	-.019	-.023	-.050	-.038	.043	.070	.127	.115
	(.026)	(.022)	(.016)	(.014)	(.070)	(.080)	(.044)	(.051)
Apr. -- June	.014	-.006	.023	.005	-.042	.016	-.059	-.011
	(.026)	(.023)	(.017)	(.015)	(.071)	(.082)	(.044)	(.051)
July -- Sep.	.049	.027	.016	.021	-.129	-.090	-.046	-.073
	(.026)	(.024)	(.017)	(.015)	(.071)	(.081)	(.045)	(.052)
Oct. -- Dec.	-.044	.002	.011	.012	.128	.004	-.021	-.031
	(.027)	(.024)	(.017)	(.015)	(.073)	(.082)	(.045)	(.052)
$Q_3 - Q_1$	.068	.049	.066	.058	-.172	-.160	-.173	-.188
	(.042)	(.037)	(.027)	(.024)	(.115)	(.131)	(.072)	(.083)
$\Sigma\|Q_i\|$	.126	.058	.101	.075	.342	.180	.254	.229
	(.060)	(.078)	(.039)	(.029)	(.164)	(.161)	(.087)	(.101)
χ^2	5.258	1.799	9.522	7.290	5.259	1.545	8.664	5.787
	[.154]	[.615]	[.023]	[.063]	[.154]	[.672]	[.034]	[.128]
N	2,464		6,185		2,464		6,185	

SOURCE: 1900 Census, with earnings imputed using 1950 Census.
NOTE: Men Born 1840-55 sample excludes individuals born or living in Massachusetts or Vermont. Men Born 1840-75 sample includes all individuals from the 1840-55 sample plus individuals born 1856-75 were unlikely to have been constrained by compulsory school attendance laws in their state (see the text for details). All models estimated with OLS. Standard errors in parentheses, *p* values in brackets. Standard errors and covariances calculated using White's (1980) method. χ^2 is for test of equality of quarter of birth dummy variables. Coefficients on quarter of birth dummy variables are restricted to sum to zero. Age is measured in quarter years. Demographic variables include lives in city with 50,000 or more inhabitants, married (1=married, living with spouse), and 8 regional dummies.

Table 5
Variation in Log Per Capita Household Income, by Child's Quarter of Birth

	Boys Born 1944-59 (1960 Census)		Boys Born 1954-69 (1970 Census)		Boys Born 1964-79 (1980 Census)	
Quarter of Birth	No Cntrls	Controls	No Cntrls	Controls	No Cntrls	Controls
Jan. -- Mar.	-.0077	-.0042	-.0056	-.0026	-.0015	.0014
	(.0029)	(.0026)	(.0013)	(.0012)	(.0014)	(.0013)
Apr. -- June	.0116	.0018	.0121	.0055	.0148	.0087
	(.0029)	(.0026)	(.0013)	(.0012)	(.0014)	(.0013)
July -- Sep.	-.0073	-.0031	-.0049	-.0025	-.0008	-.0054
	(.0028)	(.0025)	(.0013)	(.0012)	(.0013)	(.0013)
Oct. -- Dec.	.0034	.0055	-.0049	-.0003	-.0054	-.0034
	(.0028)	(.0025)	(.0013)	(.0012)	(.0013)	(.0013)
F	7.757	2.318	29.313	7.100	42.954	18.044
	[.002]	[.073]	[.000]	[.000]	[.000]	[.000]
N	280,324		1,176,484		1,191,536	

NOTE: Samples are limited to whites and blacks. All models estimated with OLS. Standard errors in parentheses, p values in brackets. The sum of the quarter of birth effects is restricted to sum to zero. F is for joint significance of quarter of birth effects. All models include a constant term. Models with controls also include race (1=black) and 8 regional dummy variables.

Figure 1

Third Quarter - First Quarter Reduced Form Effect of
Quarter of Birth on Educational Attainment and Log Weekly Earnings
for 3 Angrist and Krueger Birth Cohorts