NBER WORKING PAPERS SERIES

THE EFFECT OF AGE AT SCHOOL ENTRY ON EDUCATIONAL ATTAINMENT: AN APPLICATION OF INSTRUMENTAL VARIABLES WITH MOMENTS FROM TWO SAMPLES

Joshua D. Angrist

Alan B. Krueger

Working Paper No. 3571

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1990

The authors thank Lisa Krueger for research assistance, and Gary Chamberlain and Guido Imbens for helpful comments and discussions. Thanks also go to David Card, Whitney Newey, and two anonymous referees for their comments on earlier versions. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3571
December 1990

THE EFFECT OF AGE AT SCHOOL ENTRY ON EDUCATIONAL ATTAINMENT: AN APPLICATION OF INSTRUMENTAL VARIABLES WITH MOMENTS FROM TWO SAMPLES

ABSTRACT

This paper tests the hypothesis that compulsory school attendance laws, which typically require school attendance until a specified birthday, induce a relationship between the years of schooling and age at school entry. Variation in school start age created by children's date of birth provides a natural experiment for estimation of the effect of age at school entry. Because no large data set contains information on both age at school entry and educational attainment, we use an Instrumental Variables (IV) estimator with data derived from the 1960 and 1980 Censuses to test the age-at-entry/compulsory schooling model. In most IV applications, the two covariance matrices that form the estimator are constructed from the same sample. We use a method of moments framework to discuss IV estimators that combine moments from different data sets. In our application, quarter of birth dummies are the instrumental variables used to link the 1960 Census, from which age at school entry can be derived for one cohort of students, to the 1980 Census, which contains educational attainment for the same cohort of students. The results suggest that roughly 10 percent of students were constrained to stay in school by compulsory schooling laws.

Alan Krueger
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08544

Joshua D. Angrist
Department of Economics
Harvard University
Cambridge, MA 02138

1. Introduction

Are children better off if they start school at an earlier or later age? Although this question has long concerned researchers in a variety of disciplines (see Proctor, Black, and Feldhusen (1986) for a survey), insufficient attention has been paid to the relationship between compulsory school attendance laws and the age at school entry. This paper tests the hypothesis that compulsory school attendance laws, which typically require students to attend school until they reach their sixteenth birthday, induce a relationship between years of schooling and age at school entry. We present evidence that educational attainment is related to age at school entry because children who enter school at an older age are permitted to drop out after having completed less schooling than children who enter school at a younger age.

A simple model is outlined that shows the mechanical relationship between educational attainment and age at school entry when compulsory schooling laws are binding. The model leads to a linear relationship between age at school entry and years of completed education. If age at school entry were randomly assigned, the parameter identified by a regression of education on age at school entry would be the proportion of students who are constrained to stay in school by compulsory schooling laws. But because a nonrandom sample of children are likely to be enrolled in school at an earlier age by their parents, Ordinary Least Squares (OLS) estimation may give a biased estimate of the effect of age at school entry on educational achievement. To control for the possible endogeneity of age at entry to school, we use the fact that children born in different months of the year start school at different ages. If, as seems plausible, season of birth is uncorrelated with other determinants of education, it provides a valid instrumental variable for age at school entry.

Because a large data set that contains information on both age at school entry and educational attainment does not exist, we use an Instrumental Variables (IV) estimator that combines data derived from two independent samples. Instrumental variables estimators are a function of two sample covariance matrices. In the typical application, both sets of moments in the IV formula are estimated from the same sample of data. We discuss theoretical properties of IV estimators in which the moments underlying the estimates are derived from two independent samples. In general, Two-Sample Instrumental Variables estimators may be used whenever a set of instruments is common to two data sets but endogenous regressors and the dependent variable are included in only one or the other data sets.

We estimate the impact of compulsory schooling by combining information from the 1960 Census on the age upon entry to school of children born 1946-52 with information from the 1980 Census on the ultimate educational attainment of these same birth cohorts. Children born earlier in the year are found to enter school at an older age, and to attain less schooling. The coefficient estimates suggest that from 7 to 12 percent of students are constrained to stay in school by compulsory attendance laws. These results provide evidence on the efficacy of compulsory attendance laws, and are relevant for discussions of school start age policy.

2.1 Age at School Entry and Educational Attainment

Suppose that all children born in the same year enter school in the Fall of the year in which they turn six, and that they are required by law to stay in school until their sixteenth birthday. Suppose also that a fixed fraction, π, would like to drop out of school as soon as possible, and that this fraction is independent of quarter of birth. If students are compared on their sixteenth birthday, those who were born earlier in the year will have spent less time in school than students who were born later in the year. Therefore, assuming that a fixed fraction of students drop out of school upon attaining the legal dropout age, students born earlier in the year will, on average, have less education than students born later in the year.

Elsewhere (Angrist and Krueger 1990), we present evidence from the 1960, 1970, and 1980 Censuses showing that substantial numbers of students drop out of school around the birthday when they become eligible to leave school. Students who have just turned sixteen in states where there is an age sixteen minimum schooling requirement have a larger decline in enrollment than students who have just turned sixteen in states where there is an age seventeen or eighteen schooling requirement. Micklewright (1989) also finds evidence of excessive drop out rates around the school leaving age (16) in the United Kingdom.

Our model of the effects of age at school entry and compulsory schooling is depicted in Figure 1, which shows events related to schooling on a time line for children born in a given year. In the figure time is measured in quarters of years, and children are born in periods 1-4, indexing their quarter of birth. Children may enter school at either time t_0 or be held back to enter at t_1. Students are allowed to drop out of school at age a^*, which is reached in period $d_q = q + a^*$ for students born in q. Students who do not drop out go on to complete e quarters of schooling. Students who complete school finish in period $t_0 + e$ or in period $t_1 + e$, depending on when they started school.

The probability of entering school at t_0 is assumed to be h_q for children born in q, with the remaining fraction $(1-h_q)$ entering in t_1. The possibility for delayed entry to school is introduced because students born in different quarters of the year may be more or less likely to be held back for an additional year before starting school.

A relationship between educational attainment and quarter of birth arises because a fixed fraction (π) of students are assumed to drop out of school as soon as they are legally permitted, and because start age varies by quarter of birth. Moreover, the relationship between start age and years of schooling is linear in this framework. To show this formally, denote students' age at entry to school (measured in quarters of years) by a. Then age at entry to school conditional on quarter of birth is:

(1) $$E[a|q] = h_q(t_0 - q) + (1-h_q)(t_1 - q) = t_1 - (q + 4h_q),$$

where we have made use of the fact $t_1 - t_0 = 4$.

Now, define y as completed quarters of schooling. Then

(2) $$\begin{aligned} E[y|q] &= (1-\pi)e + \pi\{ h_q[d_q - t_0] + (1-h_q)[d_q - t_1] \} \\ &= [(1-\pi)e + \pi(a^* - t_1)] + \pi[q + 4h_q]. \end{aligned}$$

Because students are held back with a different probability (h_q) in each quarter, equation (2) shows that (like age at school entry) educational attainment does not necessarily increase or decrease linearly with quarter of birth. In fact, the relationship between education and quarter of birth need not be monotonic.

Although the relationship between education and quarter of birth is not restricted by the compulsory schooling-dropout model, the relationship between education and age at entry to school is linear. To see this note that substituting (1) into (2) gives:

$$(3) \qquad E[y|q] = [(1-\pi)e + \pi a^{*}] - \pi E[a|q].$$

Thus, a regression of the average years of education attained by each quarter-of-birth cohort on a constant and the average age of the cohort when it entered school identifies the fraction of students constrained by compulsory schooling laws (π).

It should be stressed that the conventional view in the educational psychology literature is that students who are older when they start school have higher academic achievement because of their increased maturity (e.g., DiPasquale, Moule, and Flewelling 1980). In contrast, the age at entry/compulsory schooling model predicts the opposite -- students who start school at an older age should, on average, attain less schooling because they are permitted to drop out earlier in their academic careers.

Lastly, note that estimation of (3) can be interpreted as an application of instrumental variables where the underlying micro regression equation is

$$(4) \qquad y_i = \alpha - \pi a_i + \varepsilon_i,$$

where i indexes individuals, and α is the intercept in (3). The instruments are dummy variables denoting quarter of birth. Use of a full set of mutually exclusive dummy variables as instruments is the same as grouped estimation, where the groups correspond to each cell indicated by the instruments (Friedman 1957, Angrist 1991).

Given a single data set with information on both age at school entry and educational attainment, it would be possible to tabulate OLS and IV estimates of (4). However, we are not aware of any large data set containing information on both age at school entry and completed schooling. Moreover, it seems likely that OLS estimates would be inconsistent; on average, children who start school at a younger age may do so because they show signs of above average learning potential. Estimation of (4) by IV using quarter of birth as an instrumental variable is a way to overcome bias from possible correlation between age at entry and the error term in (4). Below, we discuss IV techniques with moments from two samples.

3.1 Review of Instrumental Variables Estimation

In IV estimation, the model of interest is:

$$y_i = X_i\delta + \epsilon_i, \quad i = 1, \ldots, n$$

where X_i and ϵ_i may be correlated and δ is the q×1 vector of coefficients to be estimated. The data usually consist of a single sample of size n, assumed here to be i.n.i.d., containing observations on y_i, X_i and a set of r instrumental variables, Z_i. In the compulsory schooling model, y_i is years of education, X_i is age at school entry, and Z_i is a set of dummy variables indicating quarter of birth.

Let y, X and Z denote the data matrices of dimensions n×1, n×q and n×r, respectively. Z is assumed to have the following properties:

$$Z'\varepsilon/\sqrt{n} \sim N(0, \Omega) \text{ and } \text{plim}(Z'X/n) = \Sigma_{zx},$$

where Ω is non-singular, and Σ_{zx} is bounded and of full column rank.

The asymptotically efficient IV estimator is also the estimator that minimizes the sample analog of the moment condition $E(Z_i'\varepsilon_i) = 0$ in an optimally weighted quadratic form. The sample analog of this condition is

$$f_n(\delta) = Z'(y - X\delta)/n,$$

and $\hat{m}_n(\delta) = nf_n(\delta)'\Omega^{-1}f_n(\delta)$ is the quadratic form minimized by the estimator. The estimator is

$$\hat{\delta} = (X'Z\Omega^{-1}Z'X)^{-1}X'Z\Omega^{-1}Z'y$$

with limiting distribution

$$\sqrt{n}(\hat{\delta} - \delta) \sim N(0, [\Sigma_{zx}'\Omega^{-1}\Sigma_{zx}]^{-1}).$$

In practice, Ω is replaced by a consistent estimate, $\hat{\Omega}$. A consistent estimator for the covariance matrix of $\hat{\delta}$ is given by $(X'Z[n\hat{\Omega}]^{-1}Z'X)^{-1}$.

Note that the IV estimator may be written as a function of two sets of sample moments. The first set consists of Z'X/n, the cross-product matrix of instruments and regressors. The second set consists of Z'y/n, the cross-product or covariance matrix for instruments and the dependent variable. In fact, the IV estimator may be thought of as arising from

Generalized Least Squares (GLS) estimation of the equation

$$Z'y/n = [Z'X/n]\delta + Z'\epsilon/n.$$

The next subsection discusses the theory of IV estimation and presents an over-identification test for the situation when $Z'y/n$ is estimated from one sample and $Z'X/n$ is estimated from another sample -- a situation we call Two-Sample Instrumental Variables (TSIV).

3.2 Two-Sample Instrumental Variables

The two-sample approach to instrumental variables requires that in principle observations on y, X and Z could have been drawn from each of the populations sampled in the two data sets. The two samples are denoted by

$$S_h = \{(y_{hi}, X_{hi}, Z_{hi}); i = 1, \ldots n_h\}; h = 1, 2.$$

The data sets that the researcher has access to, however, contain only y_1 and Z_1 on the one hand, and X_2 and Z_2 on the other. Note that the instruments (Z) must be available in each data set.

A set of assumptions sufficient for two-sample estimation may be formally stated as:

Assumption A1. (i) $\lim_{n_h \to \infty} E(Z_h'X_h/n_h) = \operatorname{plim}_{n_h \to \infty} (Z_h'X_h/n_h) = \Sigma_{zx}$ and

$$\sqrt{n_h}(Z_h'X_h\delta/n_h - \Sigma_{zx}\delta) \sim N(0, \omega_h) \text{ for } h = 1, 2.$$

(ii) $(Z_1'[y_1 - X_1\delta])/\sqrt{n_1} = Z_1'\epsilon_1/\sqrt{n_1} \sim N(0, \gamma_1)$.

Now, let

$$g_n(\delta) = Z_1'y_1/n_1 - Z_2'X_2\delta/n_2.$$

and note that $Z_1'y_1 - Z_2'X_2\delta = Z_1'\epsilon_1 + (Z_1'X_1 - Z_2'X_2)\delta$. Assumption A1 therefore implies the moment restrictions motivating two sample estimation: $g_n(\delta)$ has a probability limit of zero as n_1 and $n_2 \to \infty$.

Assumption A1 is sufficient for two-sample estimation and testing. Description and application of the two-sample estimator, however, is simplified by restricting the two samples as follows:

Assumption A2. (i) Moments estimated from sample 1 are independent of moments estimated from sample 2.

(ii) Write $n_2[n_1]$ to indicate that n_2 is to be viewed as a function of n_1. Then, $\lim_{n_1 \to \infty} (n_1/n_2[n_1]) = k$ for some constant, k.

These assumptions lead to a simple, computationally feasible form for the optimal weighting matrix, and allow use of a $\sqrt{n_1}$ normalization instead of the somewhat more cumbersome normalization of moments from S_h by $\sqrt{n_h}$. In particular, given A2, $\sqrt{n_1}g_n(\delta)$ has limiting covariance $\Phi = \phi_1 + k\omega_2$, where ϕ_1 is the limiting covariance matrix of $([Z_1y_1/\sqrt{n_1}] - \Sigma_{zx}\delta)$. To see this, note that

$$g_n(\delta) = (Z_1'y_1/n_1 - \Sigma_{zx}\delta) - (\sqrt{n_1}/\sqrt{n_2})([Z_2'X_2\delta/\sqrt{n_1}\sqrt{n_2}] - [(\sqrt{n_2}/\sqrt{n_1})\Sigma_{zx}\delta])$$

so that as $n_1 \to \infty$

$$\sqrt{n_1} g_n(\delta) \sim (\sqrt{n_1}[Z_1'y_1/n_1 - \Sigma_{zx}\delta] - \sqrt{k}\, \sqrt{n_2}[Z_2'X_2\delta/n_2 - \Sigma_{zx}\delta])$$
$$\sim N(0, \phi_1 + k\omega_2).$$

The optimal TSIV estimator, $\tilde{\delta}$, minimizes a Generalized Method of Moments (GMM) quadratic form $\tilde{m}_n(\delta) = n_1 g_n(\delta)'\Phi^{-1} g_n(\delta)$. This quadratic form is minimized by

$$\tilde{\delta} = ([X_2'Z_2/n_2]\Phi^{-1}[Z_2'X_2/n_2])^{-1}[X_2'Z_2/n_2]\Phi^{-1}[Z_1'y_1/n_1].$$

Note that Φ in the formula for $\tilde{\delta}$ may be replaced by $\hat{\Phi}/n_1 = \hat{\phi}_1/n_1 + \hat{\omega}_2/n_2$, where $\hat{\phi}_1$ and $\hat{\omega}_2$ denote sample estimates. Thus, restricting the limit of n_1/n_2 in A2(ii) does not affect the estimator that is actually used.

Distribution theory for the TSIV estimator and an over-identification test statistic are presented in Lemmas 1 and 2. The proofs are straightforward and are relegated to the Appendix. We note that Arellano and Meghir (1988) have independently derived the limiting distribution of a two sample estimator using the optimal minimum distance framework outlined by Chamberlain (1982).

Lemma 1. $\sqrt{n_1}(\tilde{\delta} - \delta) \sim N(0, \Psi)$ where $\Psi = (\Sigma_{zx}'\Phi^{-1}\Sigma_{zx})^{-1}$.

Dividing Ψ by n_1, it can be shown that the variance of $\tilde{\delta}$ is consistently estimated by

$$([X_2'Z_2/n_2][(\hat{\phi}_1/n_1)+(\hat{\omega}_2/n_2)]^{-1}[Z_2'X_2/n_2])^{-1}.$$

Notice that in two-sample IV estimation, each of the cross-product matrices must be divided by the appropriate sample size. Failure to divide by sample sizes leads to an inconsistent estimate unless $\lim_{n_1 \to \infty}(n_1/n_2[n_1]) = 1$. To see this observe that

$$(X_2'Z_2\Phi^{-1}Z_2X_2)^{-1}X_2'Z_2\Phi^{-1}Z_1'y_1 =$$
$$\{ [(X_2'Z_2/n_2)\Phi^{-1}(Z_2'X_2/n_2)]^{-1}(X_2'Z_2/n_2)\Phi^{-1}(Z_1'X_1\delta/n_1)(n_1/n_2) \}$$
$$+ \{ [(X_2'Z_2/n_2)\Phi^{-1}(X_2'Z_2/n_2)]^{-1}(X_2'Z_2/n_2)\Phi^{-1}(Z_1'\epsilon_1/n_1)(n_1/n_2) \}.$$

The second term has a probability limit of zero because $Z_1'\epsilon_1/\surd n_1$ has a mean-zero limiting distribution, but the first term has probability limit equal to δk.

The GMM over-identification test statistic measures the correlation between Z and ϵ when there are more instruments than endogenous regressors, and is a specification test for the assumptions underlying IV estimation. The over-identification test statistic for TSIV is the GMM minimand evaluated at $\tilde{\delta}$. This result is presented in Lemma 2.

Lemma 2. $\tilde{m}_n(\tilde{\delta}) \sim \chi^2(r-q)$.

Observe that $\tilde{m}_n(\tilde{\delta})$ is simply the GLS quadratic form for a regression of $y_1'Z_1/n_1$ on $Z_2'X_2/n_2$ using $[\Phi/n_1]^{-1}$ as the GLS weighting matrix. Again, in practice $[\Phi/n_1]^{-1}$ may be replaced by $\hat{\phi}_1/n_1 + \hat{\omega}_2/n_2$.

4.1 Empirical Analysis

To estimate equation (4), we use data from two censuses. The sample containing age at school entry is drawn from the 1960 Census, 1% Public-Use Sample. The sample containing years of education is drawn from the 1980

Census, 5% Public-Use Sample. We try to ensure that the TSIV assumption of comparability, A2(i), is satisfied by restricting both samples to males born 1946-52 in the United States. The youngest child in the sample is 7, which is above the minimum legal age for school attendance in most states in 1960. The samples are described in more detail in the data appendix.

Age at entry to first grade can be computed from the 1960 census if it is assumed that children are not held back or advanced a grade after they enter school. Under this assumption, the formula for age at entry is

$$a_i = (A_i - 2) - [(G_i - 1)*4.0] \tag{5}$$

where A_i is age measured in quarters of years on Census Day (April 1, 1960), and G_i is the grade in which the student is currently enrolled.

Years of completed educational attainment are available in the 1980 census. A limitation of the census data is that information on completed <u>quarters</u> of schooling is not available -- the highest grade completed by the individual is reported instead. The implications of using years of completed schooling instead of quarters of completed schooling as the dependent variable in equation (3) can be examined by substituting the reduced form, $E[a|q]$, for a in the underlying latent variable model (Newey (1986). Write the unobserved completed quarters of education for individual i with birthday in q as

$$y^*_{iq} = [(1-\pi)e + \pi a^*] - \pi E[a|q] + \nu = \alpha - \pi \bar{a}_q - u \tag{6}$$

where y^*_{iq} is the latent unobserved quarters of schooling, $\bar{a}_q$ is the sample estimate of $E[a|q]$, and $\nu = \pi(E[a|q]-a) + \epsilon$. The error term, u, includes the difference between $E[a|q]$ and $\bar{a}_q$, as well as the terms in ν. Now suppose that everyone completes at least k quarters (equal to k/4 years) of education. Individuals who complete k+4 quarters of schooling complete an additional year. In this "threshold" model, completed years of schooling measured in quarters is:

(7)
$$y_{iq} = \begin{cases} k & \text{if } y^*_{iq} - k < 4 \\ k + 4 & \text{otherwise.} \end{cases}$$

Therefore,

(8)
$$\begin{aligned} E[y_{iq}|q] &= k + 4\Pr[y^*_{iq} - k - 4 > 0] \\ &= k + 4F[\bar{\alpha} - \pi\bar{a}_q], \end{aligned}$$

where $\bar{\alpha} = \alpha - k - 4$ and F is the distribution function for u.

From (8), it is apparent that the relationship between completed years of schooling and completed quarters of schooling depends partly on the shape of the distribution function, F. For example, if u is distributed uniformly on the interval [-2, +2] then equation (8) and equation (3) differ only by a constant. That is, (3) applies directly to completed years of schooling as well as to completed quarters. In other examples, the relationship between completed years of schooling and age at entry to school need not be linear. In the empirical work, we proceed on the assumption that (3) gives a good approximation to (8). In principle, if the linear model is inappropriate, the IV over-identification test statistic should provide evidence of misspecification.

The sample moments for each quarter of birth from 1946 through 1952 are reported in Table 1. The estimates of average age at entry show a sawtooth pattern, with boys born in later quarters entering the first grade at a younger age. The estimates of average educational attendance are also shown in the table and are plotted in Figure 2. Apart from trend, the graph of average educational attainment also shows a jagged pattern, with first quarter births generally having less education than the preceding fourth quarter births. We have verified this pattern in samples of men born between 1920 and 1959 in both the 1970 and 1980 censuses (Angrist and Krueger 1990). We restrict our attention here to the cohort of men born 1946-1952 because members of this cohort were in elementary school at the time of the 1960 Census.

Age at school entry is not known exactly for enrolled students in the 1960 sample, and must be estimated from information on age and grade using equation (5). Evidence that these estimates of age at school entry are a reasonable measure of true age at school entry is presented in Table 2. This table reports the pattern of age at entry in two sets of states. Most states establish a minimum age at which school entry is permitted, and there is variation in the birthday cutoff used to admit students to first grade across states. Table 2 compares the pattern of age at entry by quarter of birth between states that admit children to first grade if they have turned six by September 30 or October 1 (third quarter states), and the pattern in states that admit children to first grade if they have turned six by December 31 or January 1 (fourth quarter states). The data on permissive age of school entry was collected from state laws for 1950-58, and is reported in Appendix Table A1. We were able to classify 16 states unambiguously as third or fourth quarter cutoff states.

In states with a fourth quarter cutoff, children born in the fourth quarter will tend to be youngest among children entering the first year they are eligible to start school. In contrast, in states with a third quarter cutoff, children born in the third quarter will be youngest among children entering the first year they are eligible for school. In third quarter cutoff states, children born in the fourth quarter must wait an additional year before becoming eligible for school.

The results in Table 2 reflect the effect of school birthday cutoff policy. The F-test for the difference in the pattern of age at entry by quarter of birth shows that the two regimes lead to significantly different patterns. In states with a third quarter cutoff, children born in the third quarter are youngest when they enter school. In contrast, age at entry declines monotonically in the fourth quarter cutoff states. We interpret this difference as evidence that our proxy for age at school entry is a plausible measure of true age at entry. It should be noted, however, that in both sets of states, children born in the first quarter are oldest at school entry. The fact that students born in the fourth quarter are not actually oldest in the fourth quarter cutoff states may be due to a higher propensity of parents to enroll children born in the first through third quarters one year later than permitted. Additionally, some local school boards may deviate from the state-wide minimum entry age.

4.2 Instrumental Variables Estimates

Table 3 reports Generalized Least Squares (GLS) estimates of equation (3) for models that include different parameterizations of a cohort trend in education. The GLS weighting matrix is diagonal with diagonal elements equal to $[\hat{\phi}^2/n_1] + (.01)[\hat{\omega}^2/n_2]$, where $\hat{\phi}^2$ is the cell variance of education in the 1980 census, $\hat{\omega}^2$ is the cell variance of age at school entry in the 1960 census, and n_1 and n_2 are the corresponding sample sizes. This is the optimal weighting matrix for efficient TSIV estimation, assuming $\pi = 0.10$.

Rows 1-4 of the table show results for models with an MA(+2,-2) trend in schooling. The MA(+2,-2) trend equals $(m_{-2} + m_{-1} + m_{+1} + m_{+2}) / 4$, where m_q is the mean years of education of the cohort born q quarters before or after the current quarter. We estimate models with the MA(+2,-2) trend by including the trend term as a regressor without constraining its coefficient. We note, however, that in every specification reported in Tables 3-5, the MA(+2,-2) term enters with a coefficient that is not statistically different from one. In alternative specifications, we report results including a quadratic function of age in quarters (YOBQ, $YOBQ^2$), linear age in quarters (YOBQ), and 6 year dummies (YOB).

The remaining rows show results for models fit to cell means for year of birth/quarter of birth (YOB*QOB) interactions for each state of birth (SOB). These models all include 50 state-of-birth dummies and are equivalent to instrumental variables estimation of a model with SOB dummies and trend terms, where the instrument list includes the full set of YOB*QOB*SOB interactions. The chi-square statistics in the table are over-identification test statistics for the exclusion restrictions imposed by the TSIV estimator. Because the instruments are dummy variables, the TSIV over-identification tests measure the goodness-of-fit of the model to the cell means.

All estimates of the effect of age at entry ($-\pi$) in the table are negative. Except for those with year dummies and a linear trend, the estimates are all significant, ranging from -.06 to -.14 . Estimates based on interacting year and quarter with place of birth are more precise. Goodness-of-fit tests based on the difference in the chi-square statistics indicate that the year dummy and linear trend specifications fit the data less well than the moving average specification. But the test statistics also lead to rejection of each model in an omnibus specification test. The best fitting model, on line 5 of the table, has a chi-square value of 1467 in a distribution with 1168 degrees-of-freedom. With this number of degrees-of-freedom, classical critical values are unforgiving -- the 1% critical value in this case is around 1,275.

A problem with goodness of fit testing in this context is that there are over 400,000 observations used to calculate the sample moments. With this large a sample, even slight deviations from the null are bound to be rejected. For example, with only 200,000 observations, any of the models reported in lines 5-8 of the table would be likely to pass the omnibus goodness-of-fit test. The problem of "too many observations" in hypothesis testing has been extensively discussed in the literature (e.g., Berkson, 1938), and numerous alternatives to classical tests have been proposed. For example, in a Bayesian testing procedure such as that of Schwarz (1978), critical values are given by degrees-of-freedom times the log of the sample size. Using this criterion, each of the models in Table 3 would be found acceptable.

On the basis of the over-identification tests one might reject our assumption that season of birth is a valid instrument for start age in an education equation. Moreover, some studies claim to have uncovered a correlation between season of birth and a number of behavioral and biological outcomes. The seminal study on this topic is Huntington (1938), which argues that a genetic season of birth effect exists because genetically inferior individuals are less able to contain their sexual passions in the summer. On the other hand, Lam and Miron (1987) find evidence that the seasonal pattern of children's births is unrelated to the wealth and marital status of their parents, which suggests that quarter of birth is an exogenous variable for our purposes.

In light of the potential importance of arguments for additional season of birth effects due to psychological and other factors, results from additional specifications are reported in Table 4. The omnibus goodness-of-fit statistic for an over-identified model is asymptotically equivalent to a Wald test for the equality of alternative estimates of the same parameter (Newey and West 1987). With enough data, small differences in the estimates will lead to rejection in the omnibus specification test. But small, statistically significant differences may be of little practical importance. Table 4 therefore explores the robustness of estimates calculated under alternative exclusion restrictions. Each of the specifications reported in Table 4 includes the MA(+2, -2) term to control for cohort trends.

For reference, line 1 of Table 4 reports the results from line 5 of Table 3. The estimates in line 2 of Table 4 are from a specification that includes as regressors dummy variables for second and third quarter of birth (each interacted with place of birth), so that the excluded instruments used to identify π are only interactions with a dummy for first quarter birth. The difference between the chi-square statistics in lines 1 and 2 is 147, while the difference in degrees-of-freedom is only 102. Line 2 therefore represents a statistically significant improvement over line 1 (critical $\chi^2_{.01}(147) \approx 135$), although the difference in the estimates is small relative to sampling variance.

Line 3 of Table 4 reports results where the effect of age at entry is allowed to vary with year of birth. The estimated effect of age at entry is negative for each year of birth, but the individual coefficients are not estimated precisely enough for comparison with previous estimates to be meaningful. The last set of estimates in Table 4 allows the effect of age at entry to vary with year of birth as in line 3, and includes dummies for second and third quarter births in the equation. Here, the effect of age at entry is well determined for each year of birth. The estimate of π ranges from 0.08 for men born in 1946, to 0.124 for men born in 1952. The increase in estimated π with year of birth may reflect changing behavior, or an improvement in the quality of the measure of age at entry for younger children.

The chi-square for line 4 takes on a value of 1,285 with 1,060 degrees of freedom. This represents a substantial improvement in fit over line 1, but still exceeds classical critical values for the omnibus specification test. However, the estimates in Tables 3 and 4 appear remarkably insensitive to the details of model specification. We interpret this as providing some support for the age at entry/compulsory schooling model, and as evidence that other possible effects of age at school entry on educational attainment are small.

5. Age at Entry and Compulsory Schooling for Men with Some College

As a final test of the compulsory schooling model, we consider the impact of age at school entry on education for men with at least one year of post high school education. If students have satisfied the compulsory schooling law, our simple model predicts that age at school entry will have no effect on their educational attainment. On the other hand, if genetic or psychological factors cause a relationship between age at entry and educational achievement, one would expect to continue to find a relationship between age at entry and years of education for exempt students. Because all compulsory attendance laws in the U.S. exclude students who have graduated from high school, estimating the models in Table 3 for a sample of men who attended college is one way to remove the effect of compulsory attendance laws. In the notation of the previous section, we assume that for the sub-sample with some college education, π equals zero.

Restricting the sample to men with at least one year of post high school education may appear problematic because this conditions on the dependent variable. However, this selection rule does not invalidate our test because, under the null hypothesis that age at school entry is unrelated to academic achievement ($\pi = 0$) for this sample, the mean of schooling by quarter of birth should be constant even after conditioning on college attendance.

Table 5 presents results of a test of this implication of the age at entry-compulsory schooling model. In five of the eight specifications reported in Table 5, age at entry now has a positive effect on educational attainment, and in three of the eight specifications there is a negative relationship between school start age and educational attainment. The relationship between post high school education and age at entry is less significant than that estimated in Table 3 even though the standard errors are roughly the same magnitude. Nevertheless, the results are ambiguous because of the contrasting and statistically significant estimates on lines 4-7. But it seems clear that age at entry has a smaller effect on the education of men who attended college than in the unrestricted sample.

6. Summary and Conclusions

Some studies of academic performance have argued that students have an advantage if they start school at an older age, while others have argued that students are better off if they start at an earlier age. The outcome variable that this literature typically examines is children's achievement test scores in the primary grades. Although these studies are primarily based on small samples of observations, they have generally concluded that older school entrants fare better (DiPasquale, Moule, and Flewelling 1980; Warren, Levin and Tyler 1986). An important shortcoming of this work is that the age that children enter school is treated as an exogenous variable. Furthermore, as Lewis and Griffen (1981) point out, few of the past studies of the effect of season of birth on behavioral and biological outcomes control for age (i.e., cohort effects in a cross-section).

Our paper differs from the previous literature in that we examine the effect of children's age when they start school on their eventual years of completed schooling. The number of years of education that children attain may be a better measure of their academic success than their performance on an aptitude test at an early age. In addition, we use the exogenous variation in start age that stems from the quarter of the year a child is born, as well as school admittance age requirements, to identify the effect of start age on students' eventual years of education. Finally, we adjust for cohort trends in schooling by including a variety of trend terms.

A simple model is presented that shows that if start age affects educational attainment only because of compulsory school attendance laws, then the relationship between start age and education is linear. We present a framework for Two-Sample IV estimation, and use data from two independent samples to estimate the effect of start age on educational attainment. In contrast to most of the past literature on students' test performance, we find that older entrants tend to obtain slightly less education. The estimates indicate that roughly 10 percent of all the high school students in our sample were constrained to stay in school by compulsory schooling laws.

We find only weak evidence that start age has any effect on years of post-high school education, which leads us to conclude that the case for a an effect of school start age on academic achievement beyond the effects of compulsory schooling is modest at best. Of course, this conclusion applies only to differences in school start age that are associated with season of birth. Nonetheless, this finding should be relevant for school districts that are considering changes in school entrance policies. Finally, and perhaps most importantly, our results suggest that compulsory schooling laws are an effective means of raising educational attainment.

Appendix

Proof of Lemma 1.

$Z_1'y_1 = Z_2'X_2\delta + (Z_1'y_1 - Z_2'X_2\delta)$ so that

$$\sqrt{n_1}(\hat{\delta} - \delta)$$
$$= \sqrt{n_1}([X_2'Z_2/n_2]\hat{\Phi}^{-1}[Z_2'X_2/n_2])^{-1}[X_2'Z_2/n_2]\hat{\Phi}^{-1}([Z_1'y_1/n_1 - [Z_2'X_2\delta/n_2])$$
$$\sim (\Sigma_{zx}'\Phi^{-1}\Sigma_{zx})^{-1}\Sigma_{zx}'\Phi^{-1}\sqrt{n_1}g_n(\delta),$$

by A1(i). The result then follows from A1.

Proof of Lemma 2.

Let $P_n = [I_r - Z_2'X_2(X_2'Z_2\hat{\Phi}^{-1}Z_2'X_2)^{-1}X_2'Z_2\hat{\Phi}^{-1}]$. Then we have:

$$\lim_{n_2\to\infty} P_n = P = [I_r - \Sigma_{zx}(\Sigma_{zx}'\Phi^{-1}\Sigma_{zx})^{-1}\Sigma_{zx}'\Phi^{-1}].$$

Now, $\sqrt{n_1}g_n(\hat{\delta}) = \sqrt{n_1}\{P_n([Z_1'y_1/n_1] - [Z_2'X_2\delta/n_2]) + P_n[Z_2'X_2\delta/n_2]\}$. But $P_nZ_2'X_2 = 0$ so $\sqrt{n_1}g_n(\hat{\delta})$ is simply $\sqrt{n_1}P_ng_n(\delta)$. Let $[P_n\hat{\Phi}^{-1}P_n']^-$ be any generalized inverse. Then a chi-square statistic may be formed from $\sqrt{n_1}g_n(\hat{\delta})$ as

$$n_1g_n(\hat{\delta})'[P_n\hat{\Phi}^{-1}P_n']^-g_n(\hat{\delta})$$
$$= n_1[y_1'Z_1/n_1]P_n'(P_n\hat{\Phi}^{-1}P_n')^-P_n[Z_1'y_1/n_1],$$

where the last equality is also consequence of the fact that $P_nZ_2'X_2 = 0$. It remains to show that $P_n'(P_n\hat{\Phi}^{-1}P_n')^-P_n = P_n\Phi^{-1}P_n$. The algebra for this result follows that in Newey (1985, Proof of Proposition 2), and is therefore omitted. Finally, substituting for $P_n'(P_n\hat{\Phi}^{-1}P_n')^-P_n$ gives

$$n_1[y_1'Z_1/n_1]P_n'\hat{\Phi}^{-1}P_n[Z_1'y_1/n_1] = \tilde{m}_n(\hat{\delta}).$$

The degrees of freedom of the chi-square statistic are given by the rank of $[P\Phi^{-1}P']$. Assuming that Φ is of full rank, $[P_n\Phi^{-1}P_n']$ converges to a matrix with rank equal to the rank of P, which is equal to the number of over-identifying exclusion restrictions, r - q.

Data Appendix

The samples used from the 1960 and 1980 Censuses are described below.

1960 Census

The 1960 census data set is ICPSR (1989-90) Study number 7756:
1960 Census of Population and Housing, 1960 Public Use Sample, One-In-One Hundred File.

The sample used in our analysis includes black and white boys born in the US between 1946 and 1952, who were enrolled in school in 1960.
For comparability with the 1980 sample of black and white men described below, the sample excludes boys of "Puerto Rican stock", boys with Spanish surnames in five southwestern states, and boys for whom date of birth information or school enrollment variables were allocated.

1980 Census

The 1980 census data set is ICPSR (1989-90) Study number 8101:
1980 Census of Population and Housing Public Use 5 percent Micro Data Public Use Sample (A Sample).

The sample used in our analysis includes black and white men born in the US between 1946 and 1952, who had at least 1 year of school completed in 1980. White men in the 1980 census are non-hispanic, which is why we excluded hispanic boys from the 1960 sample. The sample also excludes men for whom sex, age, quarter of birth, race, years of schooling, weeks worked in 1979 or salary in 1979 were allocated.

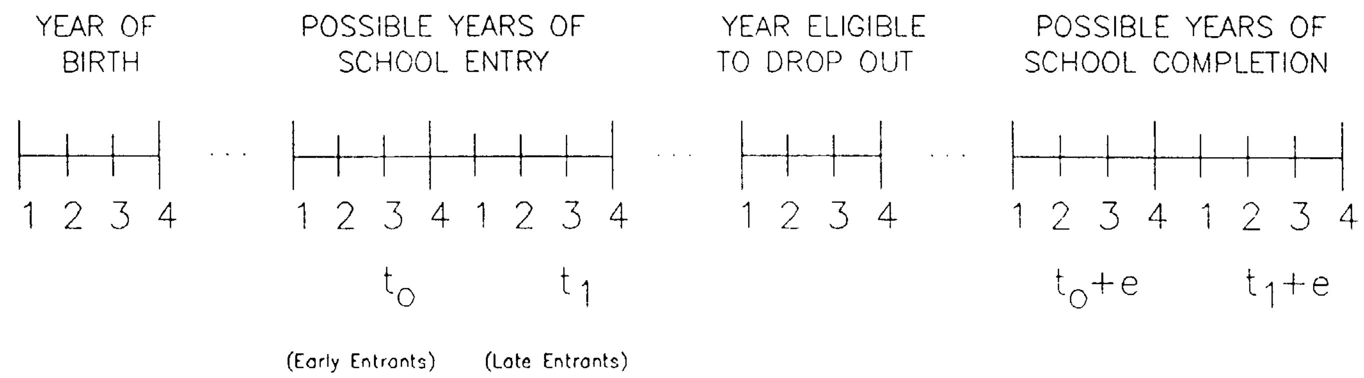
Table A1

State Laws Regarding Permissive Age at School Entry in 1955

State	1st Grade Adm. Age	Birthday Cutoff	Statute
Alabama	6	01-Oct	S:16-28-4
Arizona	6	01-Oct	S:15-821
Arkansas	6	01-Oct	S:6-18-202
California	6	01-Dec	S:480000-2
Colorado	6	01-Sep	S:22-33-104
Connecticut	6	01-Jan	S:10-15c
Delaware	6	01-Sep	S:14-203-204
Florida	6	01-Jan	S:232.01
Georgia	6	NA	-
Idaho	6	16-Oct	S:33-201
Illinois	6	01-Dec	S:10-20.12
Indiana	6	NA	S:20-8.1-3
Iowa	6	15-Sep	S:282.1-282.3
Kansas	6	01-Sep	S:72-1107
Kentucky	6	01-Oct	S:159.010
Louisiana	6	01-Dec	S:17.221.3
Maine	6	15-Oct	S:859
Maryland	6	01-Sep	S:7-101
Massachusetts	6	NA	S:76-1
Michigan	6	01-Sep	S:380.1561
Minnesota	5	NA	S:120.06
Mississippi	6	01-Jan	S:37-15-9
Missouri	6	01-Oct	S:160.051
Montana	6	10-Sep	S:20-5-101
Nebraska	6	15-Oct	S:79-444
Nevada	6	31-Dec	S:392.040
New Hampshire	6	13-Sep	S:193:1
New Jersey	6	01-Oct	S:18A:38-5
New Mexico	6	01-Jan	S:28-8-2
New York	6	01-Dec	S:1712
North Carolina	6	01-Oct	S:115C-364
North Dakota	6	31-Oct	S:15-47-02
Ohio	6	13-Sep	S:3-321.01
Oklahoma	6	01-Nov	S:1-114
Oregon	6	01-Sep	S:339.115
Pennsylvania	6	01-Feb	S:13-1304
Rhode Island	6	31-Dec	S:16-2-27,28
South Carolina	6	01-Sep	S:21-752
South Dakota	6	01-Sep	S:13-28-2
Tennessee	6	31-Dec	S:49-6-3001
Texas	6	01-Sep	S:21.031
Utah	6	02-Sep	S:53A-3-402
Vermont	6	01-Sep	T.16-S:1121
Virginia	6	30-Sep	22.1-254
Washington	6	NA	S:28A.58.190
West Virginia	6	01-Nov	S:18-2-5,18
Wisconsin	6	31-Dec	S:112,118
Wyoming	6	15-Sep	S:55 & 57

Figure 1

Time Line For A
Single Year Of Birth



<u>Notation</u>

q = quarter of birth = 1, 2, 3, or 4

t_0 = early entry date; a boy born in q enters at t_0 with probability h_q

t_1 = late entry date; a boy born in q enters a t_1 with probability $1 - h_q$

a^* = legal drop-out age in quarters

d_q = earliest drop-out date = $q + a^*$

e = educational attainment of non-drop-outs

Figure 2

Average Education by Quarter of Birth

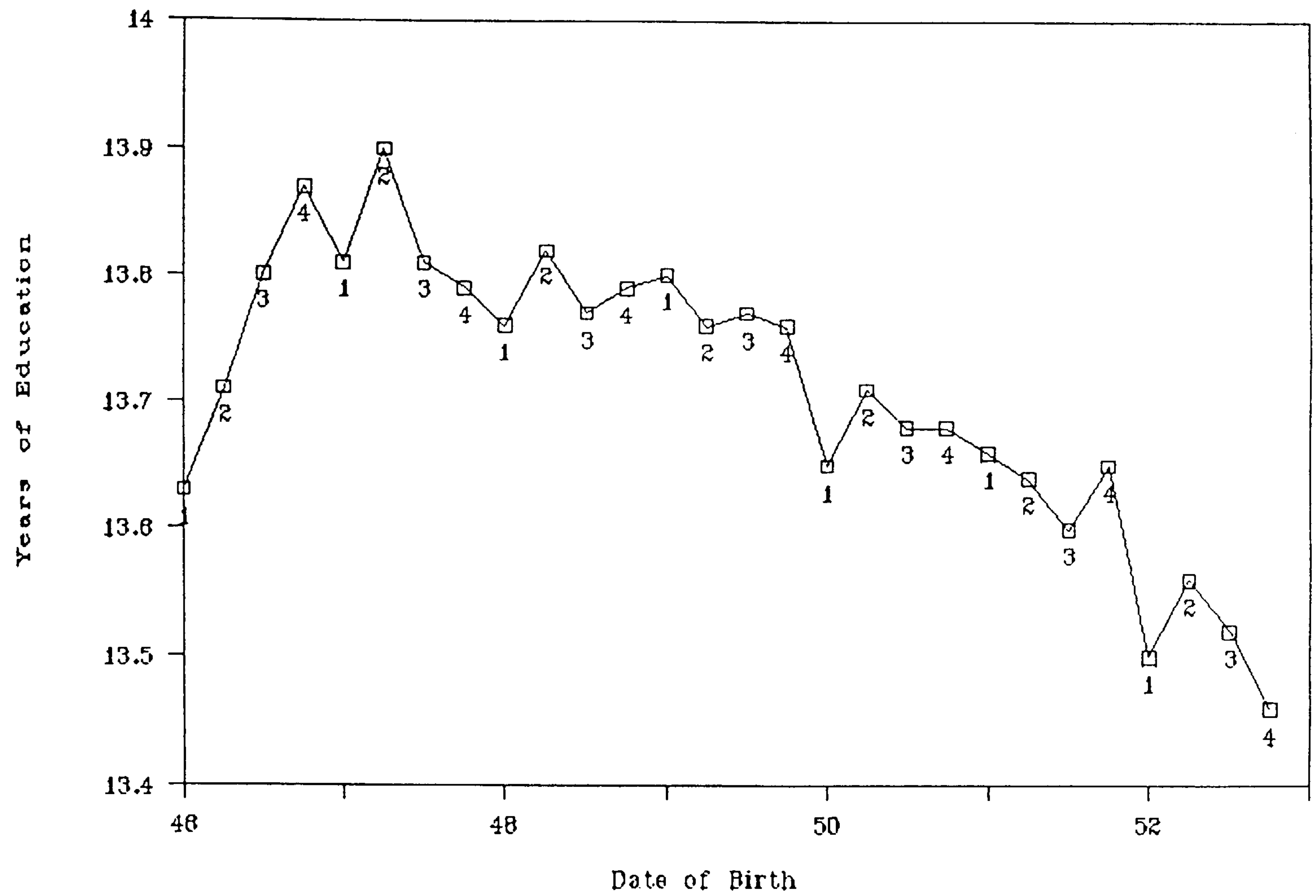


TABLE 1

EDUCATION AND AGE AT SCHOOL ENTRY

Year	Quarter	Education	Age at Entry
1946	1	13.63	6.63
	2	13.71	6.60
	3	13.80	6.38
	4	13.87	6.37
1947	1	13.81	6.64
	2	13.90	6.53
	3	13.81	6.33
	4	13.79	6.34
1948	1	13.76	6.62
	2	13.82	6.49
	3	13.77	6.33
	4	13.79	6.29
1949	1	13.80	6.61
	2	13.76	6.46
	3	13.77	6.28
	4	13.76	6.29
1950	1	13.65	6.62
	2	13.71	6.44
	3	13.68	6.24
	4	13.68	6.23
1951	1	13.66	6.54
	2	13.64	6.35
	3	13.60	6.18
	4	13.65	6.17
1952	1	13.50	6.45
	2	13.56	6.28
	3	13.52	6.08
	4	13.46	6.07

Sample:

Boys born 1946-52 in 1980 Census, 5% Public-Use Sample; US-born with at least 1 year of schooling. Sample size is 409,782.

Boys born 1946-52 in 1960 Census, 1% Public-Use Sample; US-born, enrolled in 1960. Sample for whom age at entry can be estimated has 112,033 observations.

TABLE 2

AGE AT SCHOOL ENTRY BY CUTOFF DATE

Quarter of Birth	Cutoff Date End of, Third Quarter	Fourth Quarter
1	0.319 (0.017)	0.474 (0.021)
2	0.173 (0.017)	0.277 (0.021)
3	-0.044 (0.016)	0.087 (0.020)
4	-	-
Sample Size	18,865	11,598

Notes:

Estimates are coefficients on quarter dummies in regressions of years of schooling on 3 quarter of the year born dummies and 6 year of birth dummies.

For a list of states by school entry birthday cutoff date, see Table A1. The third quarter cutoff states are: Alabama, Arizona, Arkansas, Kentucky, Missouri, New Jersey, North Carolina, and Virginia. The fourth quarter cutoff states are: Connecticut, Florida, Mississippi, Nevada, New Mexico, Rhode Island, Tennessee, and Wisconsin.

F-statistic for a test of the difference in quarter patterns by state (conditional on state main effects) is 12.6 (df = 3, prob. < .0001).

TABLE 3

EDUCATIONAL ATTAINMENT AND AGE AT ENTRY TO SCHOOL

	Instruments[a]	Regressors[b]	$-\pi$[c]	χ^2 (dof)
1.	QOB * YOB	MA(+2, -2)	-0.117 (0.070)	78.6 (21)
2.	QOB * YOB	YOBQ, $YOBQ^2$	-0.123 (0.064)	91.5 (24)
3.	QOB * YOB	YOBQ	-0.041 (0.095)	223.6 (25)
4.	QOB * YOB	YOB	-0.010 (0.064)	93.1 (20)
5.	QOB * YOB * SOB	MA(+2, -2) SOB	-0.114 (0.030)	1467 (1168)
6.	QOB * YOB * SOB	YOBQ, $YOBQ^2$ SOB	-0.138 (0.028)	1887 (1368)
7.	QOB * YOB * SOB	YOBQ, SOB	-0.092 (0.029)	2033 (1369)
8.	QOB * YOB * SOB	YOB, SOB	-0.064 (0.027)	1895 (1364)

Notes:

a. QOB denotes 3 quarter of birth dummies; YOB denotes 6 year of birth dummies; and SOB denotes 50 state of birth dummies. When denoted by an asterisk, interactions and levels of these variables are used as instruments.

b. MA(+2, -2) denotes a moving average trend term, YOBQ and $YOBQ^2$ denote a quadratic year of birth trend with year of birth measured in quarters of years; YOBQ denotes linear year of birth; YOB denotes 6 year of birth dummies; and SOB denotes 50 state of birth dummies.

c. $-\pi$ is the coefficient on the age at school entry variable.

d. Sample Information:

Boys born 1946-52 in 1980 Census, 5% Public-Use Sample; US-born with at least 1 year of schooling. Sample size is 409,782.

Boys born 1946-52 in 1960 Census, 1% Public-Use Sample; US-born, enrolled in 1960. Sample size is 112,033.

TABLE 4

EDUCATIONAL ATTAINMENT AND AGE AT ENTRY TO SCHOOL:
ALTERNATIVE EXCLUSION RESTRICTIONS

	Instruments[a]	Regressors[b]	$-\pi$[c]	χ^2 (dof)
1.	QOB * YOB * SOB	MA(+2, -2) SOB	-0.117 (0.030)	1467 (1168)
2.	QOB * YOB * SOB	MA(+2, -2) SOB QTR2*SOB QTR3*SOB	-0.165 (0.034)	1320 (1066)
3.	QOB * YOB * SOB	1946 ENTRYAGE	-0.029 (0.036)	
		1947 ENTRYAGE	-0.031 (0.036)	
		1948 ENTRYAGE	-0.036 (0.035)	
		1949 ENTRYAGE	-0.037 (0.035)	
		1950 ENTRYAGE	-0.051 (0.033)	
		1951 ENTRYAGE	-0.055 (0.032)	
		1952 ENTRYAGE	-0.070 (0.031)	
		MA(+2, -2) SOB		1434 (1162)
4.	QOB * YOB * SOB	1946 ENTRYAGE	-0.080 (0.039)	
		1947 ENTRYAGE	-0.084 (0.040)	
		1948 ENTRYAGE	-0.090 (0.039)	
		1949 ENTRYAGE	-0.087 (0.038)	
		1950 ENTRYAGE	-0.103 (0.037)	
		1951 ENTRYAGE	-0.108 (0.036)	
		1952 ENTRYAGE	-0.124 (0.035)	
		MA(+2, -2) SOB QTR2*SOB QTR3*SOB		1285 (1060)

Continued

Notes to Table 4:

a. QOB denotes 3 quarter of birth dummies; YOB denotes 6 year of birth dummies; and SOB denotes 50 state of birth dummies. Interactions of these variables are used as instruments.

b. MA(+2, -2) denotes a moving average trend term, SOB denotes 50 state of birth dummies. QTR2 and QTR3 denote dummies for second and third quarter of birth. 1946 ENTRYAGE denotes the age at school entry of men born in 1946.

c. $-\pi$ is the coefficient on the age at school entry variable.

d. Sample Information:

Boys born 1946-52 in 1980 Census, 5% Public-Use Sample; US-born with at least 1 year of schooling. Sample size is 409,782.

Boys born 1946-52 in 1960 Census, 1% Public-Use Sample; US-born, enrolled in 1960. Sample size is 112,033.

TABLE 5

EDUCATIONAL ATTAINMENT AND AGE AT ENTRY TO SCHOOL FOR THOSE WITH SOME POST-HIGH SCHOOL EDUCATION

	Instruments[a]	Regressors[b]	$-\pi$[c]	χ^2 (dof)
1.	QOB * YOB	MA(+2, -2)	0.052 (0.050)	50.1 (21)
2.	QOB * YOB	YOBQ, $YOBQ^2$	0.028 (0.038)	37.0 (24)
3.	QOB * YOB	YOBQ	0.005 (0.042)	49.8 (25)
4.	QOB * YOB	YOB	0.153 (0.038)	37.0 (20)
5.	QOB * YOB * SOB	MA(+2, -2) SOB	-0.050 (0.027)	1402 (1166)
6.	QOB * YOB * SOB	YOBQ, $YOBQ^2$ SOB	-0.055 (0.025)	1638 (1366)
7.	QOB * YOB * SOB	YOBQ, SOB	-0.064 (0.025)	1646 (1367)
8.	QOB * YOB * SOB	YOB, SOB	0.026 (0.024)	1664 (1362)

Notes:

a. QOB denotes 3 quarter of birth dummies; YOB denotes 6 year of birth dummies; and SOB denotes 50 state of birth dummies. Interactions of these variables are used as instruments.

b. MA(+2, -2) denotes a moving average trend term, YOBQ and $YOBQ^2$ denote a quadratic year of birth trend with year of birth measured in quarters of years; SOB denotes 50 state of birth dummies.

c. $-\pi$ is the coefficient on the age at school entry variable.

d. Sample Information:

Boys born 1946-52 in 1980 Census, 5% Public-Use Sample; US-born with at least 13 years of schooling. Sample size is 228,400.

Boys born 1946-52 in 1960 Census, 1% Public-Use Sample; US-born, enrolled in 1960. Sample size is 112,033.

References

Angrist, Joshua (1991), "Grouped Data Estimation and Testing in Simple Labor Supply Models," Journal of Econometrics 45, forthcoming.

__________ and Alan Krueger (1990), "Does Compulsory School Attendance Affect Schooling and Earnings?" Unpublished paper, Industrial Relations Section, Princeton University.

Arellano, Manuel and Costas Meghir (1988), "Using Complementary Data Sources: An Application to Labour Supply and Job Search," unpublished paper, University of Oxford.

Berkson, J. (1938), "Some Difficulties of Interpretation Encountered in the Application of the Chi-Square Test," Journal of the American Statistical Association 33: 526-542.

Chamberlain, Gary, (1982), "Multivariate Regression Models for Panel Data," Journal of Econometrics 18: 5-46.

DiPasquale, Glenn W., A. Moule, and R. Flewelling (1980), "The Birthdate Effect," Journal of Learning Disabilities 13(5): 234-248.

Friedman, Milton (1957), A Theory of the Consumption Function, Princeton: Princeton University Press.

Huntington, Ellsworth (1938), Season of Birth: It's Relation to Human Abilities, New York: Wiley.

Lewis, Marc, and P. Griffin (1981), "An Explanation for the Season of Birth Effect in Schizophrenia and Certain Other Diseases," Psychological Bulletin 89(3): 589-596.

Lam, David, and Jeffrey Miron (1987), "The Seasonality of Births in Human Populations," unpublished paper, Population Studies Center Research Report No. 87-114, University of Michigan.

Micklewright, John (1989), "Choice at Sixteen," Economica 56: 25-39.

Newey, Whitney K. (1985), "Generalized Method of Moments Estimation and Testing," Journal of Econometrics 29: 229-256.

__________ (1986), "Linear Instrumental Variables Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," Journal of Econometrics 32: 127-141.

__________ and Kenneth D. West (1987), "Hypothesis Testing with Efficient Method of Moments Estimation," International Economic Review 28, 777-787.

Proctor, Theron B., Kathryn Black, and John Feldhusen (1986), "Early Admission of Selected Children to Elementary School: A Review of the Research Literature," Journal of Educational Research 80(2): 70-76.

Schwarz, G. (1978), "Estimating the Dimension of a Model," The Annals of Statistics 6(2): 461-464.

Warren, Charles W., M. Levin, and Carl Tyler (1986), "Season of Birth and Academic Achievement," Educational and Psychological Research 6(2): 111-123.