The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples

JOSHUA D. ANGRIST and ALAN B. KRUEGER*

We present a model in which compulsory school attendance laws, which typically require school attendance until a specified birthday, induce a relationship between years of schooling and age at school entry. Variation in school starting age created by children's dates of birth provides a natural experiment for estimating 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 estimate and 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 compulsory attendance laws constrain roughly 10% of students to stay in school.

KEY WORDS: Compulsory attendance; Method of moments; Season of birth.

Are children better off if they start school at an earlier or later age? Although this question has long concerned researchers in various disciplines (see Proctor, Black, and Feldhusen 1986 for a survey), insufficient attention has been paid to the relationship between compulsory school attendance laws and children's ages at school entry. This article tests the hypothesis that compulsory school attendance laws, which typically require students to attend school until they reach their 16th 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 likely will 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. (See Mare [1980] for evidence that school transitions are related to family background.) 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 season of birth is uncorrelated with other determinants of education, it provides a valid instrumental variable for a regression of educational attainment on 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. IV 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 data sample. We discuss theoretical properties of IV estimators in which the moments underlying the estimates are derived from two independent samples. In general, Two-Sample IV 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 at school entry of children born 1946–1952 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 than children born later in the year. The coefficient estimates suggest that from 7% to 12% 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. In particular, our results suggest that for students who attend school beyond the compulsory schooling requirement, small differences in age at school entry have no effect on ultimate educational attainment.

* Joshua D. Angrist is Senior Lecturer, Hebrew University, Jerusalem, Israel 91905; Assistant Professor, Department of Economics, Harvard University, Cambridge, MA 02138; and Faculty Research Fellow, National Bureau of Economic Research, Cambridge, MA 02138. Alan B. Krueger is Assistant Professor, Department of Economics, Princeton University, Princeton, NJ 08544; and Faculty Research Fellow, National Bureau of Economic Research, Cambridge, MA 02138. The authors thank Lisa Krueger for research assistance and Gary Chamberlain and Guido Imbens for helpful comments and discussions. Thanks also go to two anonymous referees, an associate editor and the editor, for their comments. Financial Support was provided by the National Science Foundation (SES-9012149) and the Princeton Industrial Relations Section.

Journal of the American Statistical Association
June 1992, Vol. 87, No. 418, Applications and Case Studies

1. AGE AT SCHOOL ENTRY AND EDUCATIONAL ATTAINMENT

Suppose children 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 of children, 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 born earlier in the year will have spent less time in school than students born later in the year. Therefore, assuming that a fixed fraction of students drop out of school on attaining the legal dropout age, students born earlier in the year will, on average, attain less education than will students born later in the year.

Evidence that compulsory attendance laws actually keep children in school was presented by Ehrenberg and Marcus (1982), Lang and Kropp (1986), and Stigler (1950). But Stigler and Edwards (1978) emphasized that states passing stricter compulsory attendance laws may have had higher enrollment rates anyway. More recently, in Angrist and Krueger (1991) we used the 1960, 1970, and 1980 Censuses to show that substantial numbers of students drop out near the birthday when they become eligible to leave school. Students who have just turned 16 in states with an age 16 minimum schooling requirement have a larger decline in enrollment than students who have just turned 16 in states with an age 17 or 18 schooling requirement. Micklewright (1989) also reports excessive dropout rates near the legal 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 timeline 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.

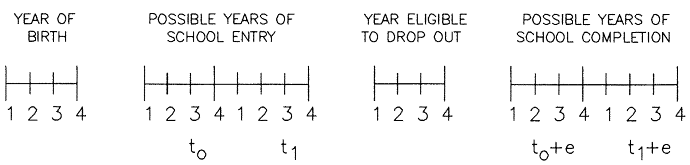

Figure 1. Notation for the Model of Age at School Entry and Educational Attainment: q = Quarter of Birth = 1, 2, 3, or 4; t_0 = Early Entry Date; a Boy Born in q Enters a t_0 With Probability h_q; t_1 = Late Entry Date; a Boy Born in q Enters a t_1 With Probability $1 - h_q$; a^* = Legal Dropout Age in Quarters; d_q = Earliest Dropout Date = $q + a^*$; e = Educational Attainment of Nondropouts.

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 to do so, and because school starting age varies by quarter of birth. 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

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

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

Now, define y as completed quarters of schooling. Then:

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

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 age at entry/compulsory schooling model, the relationship between education and age at school entry is linear. To see this, note that substituting (1) into (2) gives

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

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. (We note that in practice, this model may be only an approximation because of summer vacations—on average, children born in summer months of the third quarter complete only as much schooling as children with birthdays at the end of the school year.)

It should be stressed that the conventional view in the educational psychology literature is that students who are older when starting school attain higher academic achievement because of their increased maturity (DiPasquale, Moule, and Flewelling 1980). Consequently, some educational researchers (e.g., Uphoff and Gilmore 1985) recommend that schools modify the cutoff age for school entrance so as to exclude students born in the summer months. 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.

Implicit in our discussion of the effect of age at entry on educational outcomes is the existence of a behavioral relationship between the outcome for individual i, y_i, and age at entry, a_i. If this relationship is linear, we can write

$$y_i = \alpha + \beta a_i + \varepsilon_i,$$

where ε_i is a disturbance term representing influences on y_i other than the behavioral effect, β. The question of whether the observed correlation between y_i and a_i can be used to make inferences about β depends on whether ε_i and a_i are correlated. It seems likely that ε_i and a_i would be correlated (and therefore OLS estimates of β inconsistent), because many children who start school at a younger age may do so because they show signs of above-average learning potential.

Estimation of (3) overcomes the problem of correlation between a_i and ε_i because it is an application of IV estimation techniques to the underlying microregression

$$y_i = \alpha - \pi a_i + \varepsilon_i, \tag{4}$$

where α is the intercept in (3) and $-\pi$ is our behavioral effect. 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, in which the groups correspond to each cell indicated by the instruments. (This result is implicit in Friedman (1957, p. 35; see also Angrist [1991]). The IV estimation strategy is legitimate in this case as long as quarter-of-birth dummies are orthogonal to the error term, ε_i.

Finally, we note that given a single data set with information on both age at school entry and educational attainment, it would be possible to tabulate OLS as well as IV estimates of (4). However, we are not aware of any large data set containing information on both age at school entry and educational attainment. These variables may be derived from two separate data sets. In the following section, we discuss IV techniques with moments from two samples.

2. REVIEW OF INSTRUMENTAL VARIABLES ESTIMATION

In IV estimation, the model of interest is

$$y_i = \mathbf{X}_i\delta + \varepsilon_i, \qquad i = 1, \ldots, n,$$

where $\mathbf{X}_i$ is a $1 \times q$ row vector, ε_i a mean zero error term, and δ the $q \times 1$ vector of coefficients to be estimated. $\mathbf{X}_i$ and ε_i may be correlated. The data usually consist of a single sample of n observations, assumed here to be independent, containing observations on y_i, $\mathbf{X}_i$, and a $1 \times r$ vector of instrumental variables, $\mathbf{Z}_i$. In the compulsory schooling model, y_i is years of education, $\mathbf{X}_i$ is age at school entry and a constant, and $\mathbf{Z}_i$ is a set of dummy variables indicating quarter of birth.

Let $\mathbf{y}$, X, and Z denote the data matrices of dimensions $n \times 1$, $n \times q$, and $n \times r$. Z is assumed to have the following properties:

$$Z'\boldsymbol{\varepsilon}/\sqrt{n} \sim N(0, \Omega) \quad \text{and} \quad \operatorname*{plim}_{n\to\infty}(Z'X/n) = \Sigma_{zx},$$

where $\sim N$ denotes asymptotic normality, Ω is non-singular, and Σ_{zx} is bounded and of full column rank. Under our sampling conditions, a central limit theorem (White 1982) implies that $\Omega = \text{plim}_{n\to\infty} \sum_i \mathbf{Z}_i'\mathbf{Z}_i\varepsilon_i^2/n$.

The asymptotically efficient IV estimator minimizes the sample analog of the moment condition $E(\mathbf{Z}_i'\varepsilon_i) = 0$ in an optimally weighted quadratic form. The sample analog of this condition is $f_n(\delta) = Z'(\mathbf{y} - X\delta)/n$, and $\hat{m}_n(\delta) = nf_n(\delta)'\Omega^{-1}f_n(\delta)$ is the optimally weighted quadratic form. The resulting estimator is $\hat{\delta} = (X'Z\Omega^{-1}Z'X)^{-1}X'Z\Omega^{-1}Z'\mathbf{y}$, with limiting distribution $\sqrt{n}(\hat{\delta} - \delta) \sim N(\mathbf{0}, [\Sigma_{zx}'\Omega^{-1}\Sigma_{zx}]^{-1})$. In practice, Ω is replaced by a consistent estimate, $\hat{\Omega}$. A consistent estimator for the covariance matrix of $\hat{\delta}$ is $(X'Z[n\hat{\Omega}]^{-1}Z'X)^{-1}$. A consistent estimator of Ω for the single sample case is $\sum_i \mathbf{Z}_i'\mathbf{Z}_i\hat{\varepsilon}_i^2/n$, where $\hat{\varepsilon}_i$ is an estimate of ε_i obtained using the conventional two-stage least squares (TSLS) estimate of δ.

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 $Z'\mathbf{y}/n$, 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 $Z'\mathbf{y}/n = [Z'X/n]\delta + Z'\boldsymbol{\varepsilon}/n$. The next subsection discusses the theory of IV estimation and presents an overidentification test for the situation in which $Z'\mathbf{y}/n$ is computed using one sample and $Z'X/n$ is computed using another sample.

2.1 Two-Sample Instrumental Variables (TSIV)

We develop the two-sample approach to instrumental variables estimation by imagining that, in principle, observations on $\mathbf{y}$, X, and Z *could* have been drawn from each of the populations sampled in the two data sets. The two samples are denoted by

$$S_h = \{(y_{hi}, \mathbf{X}_{hi}, \mathbf{Z}_{hi}); i = 1, \ldots, n_h\}; \qquad h = 1, 2.$$

The data sets to which the researcher has access, however, contain only $\mathbf{y}_1$ 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. (1) $\text{plim}_{n_h\to\infty}(Z_h'X_h/n_h) = \Sigma_{zx}$, and $\sqrt{n_h}(Z_h'X_h\delta/n_h - \Sigma_{zx}\delta) \sim N(\mathbf{0}, \omega_h)$ for $h = 1, 2$. (2) $(Z_1'[\mathbf{y}_1 - X_1\delta])/\sqrt{n_1} \equiv Z_1'\boldsymbol{\varepsilon}_1/\sqrt{n_1} \sim N(\mathbf{0}, \gamma_1)$.

Now, let $g_n(\delta) = Z_1'\mathbf{y}_1/n_1 - Z_2'X_2\delta/n_2$ and note that $g_n(\delta) = Z_1'\boldsymbol{\varepsilon}_1/n_1 + (Z_1'X_1/n_1 - Z_2'X_2/n_2)\boldsymbol{\delta}$. Assumption A1 therefore implies the moment restrictions motivating two-sample estimation: $g_n(\delta)$ has probability limit 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. (1) Moments estimated from sample 1 are independent of moments estimated from sample 2. (2) 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 matrix $\Phi = \phi_1 + k\omega_2$, where ϕ_1 is the limiting covariance matrix of $([Z_1'\mathbf{y}_1/\sqrt{n_1}] - \Sigma_{zx}\delta)$. To see this, note that

$$g_n(\delta) = \{Z_1'\mathbf{y}_1/n_1 - \Sigma_{zx}\delta\} - (\sqrt{n_1}/\sqrt{n_2})\{[Z_2'X_2\delta/\sqrt{n_1}\sqrt{n_2}] - [(\sqrt{n_2}/\sqrt{n_1})\Sigma_{zx}\delta]\}$$

so that as $n_1 \to \infty$

$$\sqrt{n_1}g_n(\delta) \sim \{\sqrt{n_1}[Z_1'\mathbf{y}_1/n_1 - \Sigma_{zx}\delta] - \sqrt{k}\sqrt{n_2}[Z_2'X_2\delta/n_2 - \Sigma_{zx}\delta]\} \sim N(0, \phi_1 + k\omega_2).$$

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

$$\tilde{\delta} = ([X_2'Z_2/n_2]\Phi^{-1}[Z_2'X_2/n_2])^{-1} \times [X_2'Z_2/n_2]\Phi^{-1}[Z_1'\mathbf{y}_1/n_1].$$

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 consistent estimates. Thus, restricting the limit of n_1/n_2 in A2 (2) does not affect the estimator actually used.

Distribution theory for the TSIV estimator and an overidentification test statistic are presented in Lemmas 1 and 2. The proofs are straightforward and are presented in the Appendix.

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

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

The GMM overidentification test statistic measures the correlation between $\mathbf{Z}_i$ and ε_i when there are more instruments than endogenous regressors; this provides a specification test for the assumptions underlying IV estimation. The overidentification 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 $\mathbf{y}_1'Z_1/n_1$ on $Z_2'X_2/n_2$, using $[\hat{\Phi}/n_1]^{-1}$ as the GLS weighting matrix. Again, in practice $[\hat{\Phi}/n_1]^{-1}$ may be replaced by $\hat{\phi}_1/n_1 + \hat{\omega}_2/n_2$.

Finally, we note that the use of two samples to estimate regression coefficients dates back at least to Durbin (1953), who shows how to combine a preexisting coefficient estimate with OLS estimates from a new sample. Maddala (1971) addressed a similar problem in a likelihood framework. Arellano and Meghir (1988) independently derived the limiting distribution of a two-sample estimator in an optimal minimum distance framework applied to reduced form parameters (as in Chamberlain [1982]). The main advantage of the approach outlined here is that working directly with orthogonality conditions that are linear in parameters bypasses the need to impose nonlinear restrictions. TSIV coefficient estimates and standard errors can therefore be produced by conventional regression software.

3. 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, from the 1980 Census, 5% Public-Use Sample. We try to ensure that the TSIV assumption of comparability, A2 (1), is satisfied by restricting both samples to males born 1946–1952 in the United States. The youngest child in the 1960 sample is 7, which is above the minimum legal age for school attendance in most states. 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 entering school. Under this assumption, the formula for age at entry is

$$a_i = (A_i - 2) - [(G_i - 1) \times 4], \quad (5)$$

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 quarters of schooling is not available—the highest grade completed 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 conditional expectation $E[a|q]$ for a_i in Equation (4):

$$y_i = \alpha - \pi E[a|q_i] + \nu_i, \quad (6)$$

where $\nu_i = \pi(E[a|q_i] - a_i) + \varepsilon_i$. For illustration, suppose that everyone completes at least c quarters of education and that compulsory attendance laws can compel at most a single additional completed year of schooling. Individuals who complete at least $c + 4$ quarters of schooling complete the additional year. In this model, completed years of schooling measured in quarters (y_i^*) is

$$y_i^* = c \quad \text{if } y_i < c + 4$$
$$= c + 4 \quad \text{otherwise.} \quad (7)$$

Therefore,

$$E[y_i^*|q_i] = c + 4 - 4\Pr[y_i < c + 4|q_i] = c + 4 - 4F[c + 4 - \alpha + \pi E(a|q_i)] \quad (8)$$

where F is the distribution function for ν_i and ν_i is assumed independent of q_i.

From Equation (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 ν_i is distributed uniformly on the interval $[-2, +2]$, then Equation (8) and Equation (3) are the same. That is, (3) applies directly to completed

Table 1. Years of Education Attained and Age at School Entry

Year	Quarter	Years of education attained	Age at school 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

NOTE: Boys born 1946–1952 in 1980 Census, 5% Public-Use Sample; U.S.-born with at least 1 year of schooling. Sample size is 409,782. Boys born 1946–1952 in 1960 Census, 1% Public-Use Sample; U.S.-born, enrolled in 1960. Sample for whom age at entry can be estimated has 112,033 observations.

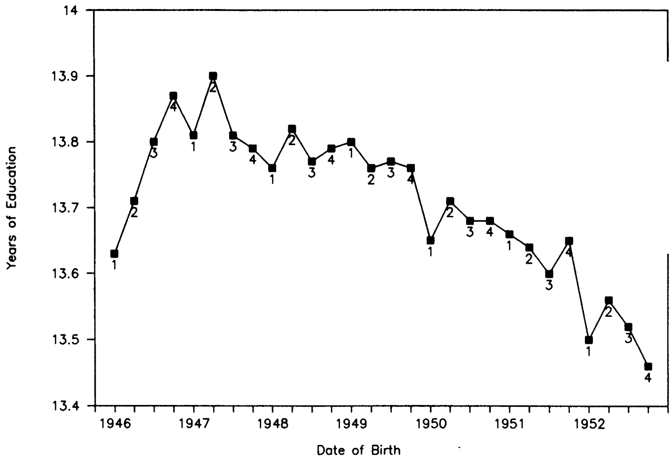

Figure 2. Average Education by Quarter of Birth for Men Born in the U.S. Between 1946–1952 With at Least One Year of Schooling. Data from the 1980 Census, 5% Public-Use Sample. Quarter of birth is indicated below each point in the figure.

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 our empirical work, we proceeded on the assumption that (3) gives a good approximation to (8). In principle, if the linear model is inappropriate, the IV overidentification 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 saw-tooth pattern, with boys born in later quarters entering the first grade at a younger age. The estimates of average educational attainment 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 children born in the first quarter generally attaining less education than children born in the preceding fourth quarter. We have verified this pattern in samples of men born between 1920 and 1959 in both the 1970 and 1980 Censuses (Angrist and Krueger 1991). 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 patterns of age at entry in two sets of states. Most states mandate a minimum age for school entry, but there is variation among states in the birthday cutoff for admitting students to first grade. Table 2 compares the patterns of age at school entry by quarter of birth in states that admit children to first grade who turn 6 by September 30 or October 1 (third-quarter-cutoff states) and states that admit children to first grade who turn 6 by December 31 or January 1 (fourth-quarter-cutoff states). The data on permitted age of school entry were collected from state laws for 1950–1958 and are reported in Appendix Table B1 on page 336. We were able to classify 16 states unambiguously as either third- or fourth-quarter-cutoff states.

In fourth-quarter-cutoff states, children born in the fourth quarter will tend to be the youngest of the children entering

Table 2. Age at School Entry by Admitting Cutoff Date

Quarter of birth	Admitting cutoff date: End of third quarter	Admitting cutoff date: End of fourth quarter
1	.319 (.017)	.474 (.021)
2	.173 (.017)	.277 (.021)
3	−.044 (.016)	.087 (.020)
4	—	—
Sample size	18,865	11,598

NOTE: Estimates are coefficients on quarter-of-birth dummies in regressions of years of schooling on three quarter-of-birth dummies and six year-of-birth dummies. For a list of states by school-entry birthday-cutoff date, see Table B.1. 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).

school the first year they are eligible. In contrast, in third-quarter-cutoff states, children born in the third quarter will be among the youngest children entering in the first year of eligibility. In third-quarter-cutoff states, children born in the fourth quarter must wait an additional year before becoming eligible for school.

The results presented 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 third-quarter-cutoff states, 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 the time of school entry. The fact that, in the third-quarter-cutoff states, students born in the fourth quarter are not actually the oldest 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 statewide minimum entry ages.

3.1 Instrumental Variables Estimates

Table 3 reports 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 $\pi = .10$.

Row 1 of the table shows 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_s is the mean years of education of the cohort born s 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 (and also in Table 4), the MA(+2, −2) term enters with a coefficient not statistically different from 1. Alternative specifications reported in rows 2–4 of Table 3 include as trend terms a quadratic function of age in quarters (YOBQ, YOBQ^2), linear age in quarters (YOBQ), and 6 year-of-birth 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 SOB dummies and are equivalent to IV 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-squared statistics in the table are overidentification test statistics for the exclusion restrictions imposed by the TSIV estimator. Because the instruments are dummy variables, the TSIV overidentification tests measure the goodness of fit of the model to the cell means.

Table 3. Educational Attainment and Age at Entry to School

Instruments[a]	*Regressors*[b]	$-\pi$[c]	χ^2 *(dof)*
QOB∗YOB	MA(+2, −2)	−.117 (.070)	78.6 (21)
QOB∗YOB	YOBQ, YOBQ^2	−.123 (.064)	91.5 (24)
QOB∗YOB	YOBQ	−.041 (.095)	223.6 (25)
QOB∗YOB	YOB	−.010 (.064)	93.1 (20)
QOB∗YOB∗SOB	MA(+2, −2) SOB	−.114 (.030)	1467 (1168)
QOB∗YOB∗SOB	YOBQ, YOBQ^2 SOB	−.138 (.028)	1887 (1368)
QOB∗YOB∗SOB	YOBQ, SOB	−.092 (.029)	2033 (1369)
QOB∗YOB∗SOB	YOB, SOB	−.064 (.027)	1895 (1364)

NOTE: Boys born 1946–1952 in 1980 Census, 5% Public-Use Sample; U.S.-born with at least one year of schooling. Sample size is 409,782. Boys born 1946–1952 in 1960 Census, 1% Public-Use Sample; U.S.-born, enrolled in 1960. Sample size is 112,033.

[a] QOB denotes three quarter-of-birth dummies; YOB denotes six year-of-birth dummies; and SOB, 50 state-of-birth dummies. Interactions and levels of these variables are used as instruments.

[b] MA(+2, −2) denotes a moving average trend term. YOBQ, YOBQ^2 denote a quadratic year-of-birth trend with year of birth measured in quarters of years. YOBQ denotes linear year of birth in quarters.

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

All estimates of the effect of age at school 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-squared 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-squared value of 1,467 in a distribution with 1,168 df. 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 more than 400,000 observations are 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 likely pass the omnibus goodness-of-fit test. The problem of "too many observations" in hypothesis testing is an old one in the statistics literature (Berkson 1938), and numerous alternatives to classical tests have been proposed. For example, in a Bayesian testing procedure such as that proposed by 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 overidentification tests, one might reject our assumption that season of birth is a valid instrument for school entry 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 was conducted by Huntington (1938), who argues that a genetic season-of-birth effect exists because genetically inferior individuals are less able to contain their sexual passions in the summer. A modern variation on this theory is found in a study by Warren and Tyler (1979). On the other hand, Lam and Miron (1987) found that the seasonal pattern of children's births is unrelated to the wealth and marital status of their parents, which suggests that birth quarter is an exogenous variable for our purposes.

In light of the potential importance of arguments for additional season-of-birth effects on education due to psychological or other factors, results from additional specifications are reported in Table 4. The omnibus goodness-of-fit statistic for an overidentified model is asymptotically equivalent to a Wald test for the equality of alternative estimates of the same parameter (Newey and West 1987). With sufficient 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 births (each interacted with place of birth), so that the excluded instruments used to identify π are interactions only with a dummy for first-quarter births. The difference between the chi-squared 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 to enable meaningful comparison with previous estimates. The last set of estimates in Table 4 allows the effect of age at entry to vary with year of birth, as in row 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 .08 for men born in 1946 to .124 for men born in 1952. The increase in estimated π with year of birth may reflect changing behavior or more accurate measurement of age at entry for younger children.

The chi-squared statistic in row 4 has a value of 1,285 with 1,060 df. This represents a substantial improvement in fit over row 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.

Table 4. Educational Attainment and Age at Entry to School: Alternative Exclusion Restrictions

Instruments[a]	*Regressors*[b]	$-\pi$[c]	χ^2 *(dof)*
QOB∗YOB∗SOB	MA(+2, −2)	−.117	1,467 (1,168)
	SOB	(.030)	
QOB∗YOB∗SOB	MA(+2, −2)	−.165	1,320 (1,066)
	SOB	(.034)	
	QTR2∗SOB		
	QTR3∗SOB		
QOB∗YOB∗SOB	1946 entry age	−.029	
		(.036)	
	1947 entry age	−.031	
		(.036)	
	1948 entry age	−.036	
		(.035)	
	1949 entry age	−.037	
		(.035)	
	1950 entry age	−.051	
		(.033)	
	1951 entry age	−.055	
		(.032)	
	1952 entry age	−.070	
		(.031)	
	MA(+2, −2)		
	SOB		1,434 (1,162)
QOB∗YOB∗SOB	1946 entry age	−.080	
		(.039)	
	1947 entry age	−.084	
		(.040)	
	1948 entry age	−.090	
		(.039)	
	1949 entry age	−.087	
		(.038)	
	1950 entry age	−.103	
		(.037)	
	1951 entry age	−.108	
		(.036)	
	1952 entry age	−.124	
		(.035)	
	MA(+2, −2)		
	SOB		
	QTR2∗SOB		
	QTR3∗SOB		1,285 (1,060)

NOTE: Boys born 1946–1952 in 1980 Census, 5% Public-Use Sample; U.S.-born with at least one year of schooling. Sample size is 409,782. Boys born 1946–1952 in 1960 Census, 1% Public-Use Sample; U.S.-born, enrolled in 1960. Sample size is 112,033.

[a] QOB denotes three quarter-of-birth dummies; YOB denotes six year-of-birth dummies; and SOB, 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 Entry Age denotes the age at school entry of men born in 1946.

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

As a final test of the age-at-entry/compulsory-schooling model, we considered the impact of age at school entry after restricting our 1980 Census sample to men with at least one year of post–high school education. Students with post-high school education have satisfied the compulsory schooling requirement, so our simple model predicts that their age at school entry has no effect on their educational attainment. In the notation of the previous section, π equals zero for this sample. On the other hand, if genetic or psychological factors cause a relationship between age at entry and educational achievement, one would expect to find a relationship between age at entry and years of education for exempt students.

Results of estimation for men with some post–high school education, reported in our working paper (Angrist and

Krueger 1990), showed little evidence of a relationship between age at entry and education after controlling for cohort trends. A related result is reported in Angrist and Krueger (1991), which showed little evidence of season-of-birth effects on completed education for large samples of college-educated men in both the 1970 and 1980 Censuses. We interpret these findings as generally supportive of the age-at-entry/compulsory-schooling model.

4. SUMMARY AND CONCLUSIONS

In various studies of academic performance, some authors have argued that students gain an advantage by starting school at an older age; others have argued that students are better served by starting at a younger age. Typically, the outcome variable examined in this literature is children's achievement test scores in the primary grades. Although these studies are based primarily on small samples of observations, they generally conclude that older school entrants fare better (DiPasquale, Moule, and Flewelling 1980; Warren, Levin, and Tyler 1986). As pointed out by Gredler (1980), however, an important shortcoming of this work is that the age that children enter school is treated as an exogenous variable, and other age effects are ignored.

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

A simple model is presented showing that, if starting age affects educational attainment only because of compulsory school attendance laws, then the relationship between starting 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 starting age on educational attainment. Our results indicate that older entrants tend to attain slightly less education, and that roughly 10% of men born between 1946–1952 were constrained to stay in school by compulsory schooling laws.

In addition to providing evidence on the efficacy of compulsory attendance laws, our results lead us to conclude that the case for an effect of school starting age on educational attainment beyond the effects of compulsory schooling is modest at best. Of course, this conclusion applies only to differences in school starting age associated with season of birth. Nonetheless, this finding should be relevant for school districts considering changes in school entrance policies and also for parents deciding when to enroll their children in school.

APPENDIX A: PROOFS

Proof of Lemma 1. Note that

$$Z'_1\mathbf{y}_1 \equiv Z'_2X_2\delta + (Z'_1\mathbf{y}_1 - Z'_2X_2\delta)$$

so that

$$\sqrt{n_1}(\tilde{\delta} - \delta) = \sqrt{n_1}([X'_2Z_2/n_2]\hat{\Phi}^{-1}[Z'_2X_2/n_2])^{-1} \times [X'_2Z_2/n_2]\hat{\Phi}^{-1}([Z'_1\mathbf{y}_1/n_1 - [Z'_2X_2\delta/n_2]) \sim (\Sigma'_{zx}\Phi^{-1}\Sigma_{zx})^{-1}\Sigma'_{zx}\Phi^{-1}\sqrt{n_1}g_n(\delta).$$

The result then follows from A1.

Proof of Lemma 2. Let

$$P_n = [I_r - Z'_2X_2(X'_2Z_2\hat{\Phi}^{-1}Z'_2X_2)^{-1}X'_2Z_2\hat{\Phi}^{-1}].$$

Then we have:

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

Now, $\sqrt{n_1}g_n(\tilde{\delta}) = \sqrt{n_1}\{P_n([Z'_1\mathbf{y}_1/n_1] - [Z'_2\delta/n_2]) + P_n[Z'_2X_2\delta/n_2]\}$. But $P_nZ'_2X_2 = 0$ so $\sqrt{n_1}g_n(\tilde{\delta}) = \sqrt{n_1}P_ng_n(\delta)$. Let $[P_n\hat{\Phi}^{-1}P'_n]^-$ be any generalized inverse. Then a chi-square statistic may be formed from $\sqrt{n_1}g_n(\tilde{\delta})$ as

$$n_1g_n(\tilde{\delta})'[P_n\hat{\Phi}^{-1}P'_n]^-g_n(\tilde{\delta}) = n_1[\mathbf{y}'_1Z_1/n_1]P'_n(P_n\hat{\Phi}^{-1}P'_n)^-P_n[Z'_1\mathbf{y}_1/n_1],$$

where the last equality is also consequence of the fact that $P_nZ'_2X_2 = 0$. 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

$$n_1[\mathbf{y}'_1Z_1/n_1]P'_n\hat{\Phi}^{-1}P_n[Z'_1\mathbf{y}_1/n_1] = \tilde{m}_n(\tilde{\delta}).$$

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

APPENDIX B: DATA

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

1960 Census

The 1960 Census data set is ICPSR (1989–1990) Study No. 7756: Census of Population and Housing, 1960 Public-Use Sample: One-In-One Hundred sample.

The sample used in our analysis includes Black and White boys born in the United States 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 (1989–1990) Study No. 8101: "Census of Population and Housing, 1980 [UNITED STATES]: Public Use Microdata Sample (A Sample): 5-percent sample."

The sample used in our analysis includes black and white men born in the United Stats between 1946 and 1952, who had at least 1 year of school completed in 1980. The definition of white men in the 1980 Census excludes Hispanics, 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, and weeks worked in 1979 or salary in 1979 were allocated.

Table B.1. State Laws Regarding Minimum Age for School Entry in 1955

State	First grade entry 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

[*Received March 1990. Revised October 1991.*]

REFERENCES

Angrist, J. D. (1991), "Grouped Data Estimation and Testing in Simple Labor Supply Models," *Journal of Econometrics,* 47, 243-266.

Angrist, J. D., and Krueger, A. B. (1990), "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples," Working Paper 3571, National Bureau of Economic Research, Cambridge, MA.

——— (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics,* 106, 979–1014.

Arellano, M., and Meghir, C. (1988), "Using Complementary Data Sources: An Application to Labour Supply and Job Search," unpublished paper, Institute of Economics and Statistics, 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, G. (1982), "Multivariate Regression Models for Panel Data," *Journal of Econometrics,* 18, 5–46.

DiPasquale, G. W., Moule, A., and Flewelling, R. (1980), "The Birthdate Effect," *Journal of Learning Disabilities,* 13, 234–248.

Durbin, J. (1953), "A Note on Regression When There is Extraneous Information About One of the Coefficients," *Journal of the American Statistical Association,* 48, 799–808.

Edwards, L. (1978), "An Empirical Analysis of Compulsory Schooling Legislation, 1940–1960," *Journal of Law and Economics,* 21, 203–222.

Ehrenberg, R., and Marcus, A. (1982), "Minimum Wages and Teenagers' Enrollment-Employment Outcomes: A Multinomial Logic Model," *Journal of Human Resources,* 27, 39–58.

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

Gredler, G. R. (1980), "The Birthdate Effect: Fact or Artifact?" *Journal of Learning Disabilities,* 13, 9–12.

Huntington, E. (1938), *Season of Birth: Its Relation to Human Abilities,* New York: John Wiley.

Lam, D., and Miron, J. (1987), "The Seasonality of Births in Human Populations," Research Report 87-114, University of Michigan, Population Studies Center.

Lang, K., and Kropp, D. (1986), "Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws," *Quarterly Journal of Economics,* 101, 609–624.

Maddala, G. S. (1971), "The Likelihood Approach to Pooling Cross-Section and Time Series Data," *Econometrica,* 39, 939–953.

Mare, R. D. (1980), "Social Background and School Continuation Decisions," *Journal of the American Statistical Association,* 75, 295–305.

Micklewright, J. (1989), "Choice at Sixteen," *Economica,* 56, 25–39.

Newey, W. K. (1985), "Generalized Method of Moments Estimation and Testing," *Journal of Econometrics,* 29, 229–256.

Newey, W. K., and West, K. D. (1987), "Hypothesis Testing With Efficient Method of Moments Estimation," *International Economic Review,* 28, 777–787.

Proctor, T. B., Black, K., and Feldhusen, J. (1986), "Early Admission of Selected Children to Elementary School: A Review of the Research Literature," *Journal of Educational Research,* 80, 70–76.

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

Stigler, G. J. (1950), "Employment and Compensation in Education," Occasional Paper 33, National Bureau of Economic Research, Cambridge, Massachusetts.

Uphoff, J. K., and Gilmore, J. (1985), "Pupil Age at School Entrance—How Many are Ready for Success?" *Educational Leadership,* 43, 86–70.

Warren, C. W., and Tyler, C. (1979), "Social Status and Season of Birth: A Study of a Metropolitan Area in the Southeastern United States," *Social Biology,* 26, 275–288.

Warren, C. W., Levin, M., and Tyler, C. (1986), "Season of Birth and Academic Achievement," *Educational and Psychological Research,* 6, 111–123.

White, H. (1982), "Instrumental Variables Regression With Independent Observations," *Econometrica,* 50, 482–499.