SCHOOLING AND THE AFQT:
EVIDENCE FROM SCHOOL ENTRY LAWS

Elizabeth U. Cascio
Ethan G. Lewis

Schooling and the AFQT: Evidence from School Entry Laws
Elizabeth U. Cascio and Ethan G. Lewis
NBER Working Paper No. 11113
February 2005
JEL No. I20, J24, J15

ABSTRACT

Is the Armed Forces Qualifying Test (AFQT) a measure of achievement or ability? The answer to this question is critical for drawing inferences from studies in which it is employed. In this paper, we test for a relationship between schooling and AFQT performance in the NLSY 79 by comparing test-takers with birthdays near state cutoff dates for school entry. We instrument for schooling at the test date with academic cohort – the year in which an individual should have entered first grade – in a model that allows age at the test date to have a direct effect on AFQT performance. This identification strategy reveals large impacts of schooling on the AFQT performance of racial minorities, providing support for the hypothesis that the AFQT measures school achievement.

Elizabeth U. Cascio
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616-8578
and NBER
eucascio@ucdavis.edu

Ethan G. Lewis
Research Department
Federal Reserve Bank of Philadelphia
Ten Independence Mall
Philadelphia, PA 19106
ethan.g.lewis@phil.frb.org

I.   Introduction

The relative importance of "nature" and "nurture" to performance on standardized tests has long been a matter of academic debate. This debate was given new life in the mid-1990s with publication of *The Bell Curve*, where Herrnstein and Murray (1994) argued that a variety of social and economic outcomes could be predicted with performance on a single cognitive test—the Armed Forces Qualifying Test, or AFQT. While this finding alone was perhaps surprising, more controversial was the authors' interpretation of the AFQT as a measure of innate ability, rather than as a measure of achievement or learned skill. Given this interpretation, their analysis implied that investments in human capital by individual choice or government intervention could do little to break the intergenerational transmission of socioeconomic status (Heckman, 1995).

Understanding whether the AFQT is a measure of ability or achievement is clearly critical for drawing inferences from studies in which it is employed. Nonetheless, there does not appear to be a consensus of interpretation in the myriad of other studies in which the AFQT—available for most respondents in the National Longitudinal Survey of Youth 1979 (NLSY 79)—has been used as a proxy for unobservable skill.[1] In part, this lack of consensus can be attributed to an identification problem. Testing the ability interpretation of the AFQT requires examining the relationship between test performance and measurable forms of human capital investment, such as completed schooling. However, schooling could itself be related to innate ability, so that individuals with higher ability will have both scored better on the test and accumulated more years of schooling by the time it was administered. Simple correlations, such as that presented in Fischer, *et al.* (1996) or Bock and Moore (1986), may therefore overstate the effect of schooling, biasing researchers toward rejecting the ability interpretation of the test.

---

[1] Many of these studies have been catalogued at http://www.chrr.ohio-state.edu/nls-bib/.

Testing the ability interpretation of the AFQT is further complicated by the strong correlation between completed schooling and age. Developmental psychologists have long thought of chronological age as a metric for biological maturation and cumulative life experience, including informal investments in human capital (e.g., being read to at home during childhood).[2] Therefore, both age and formal investments in human capital, like schooling, might independently affect how well an individual performs on tests like the AFQT. Separating the effects of schooling and age is particularly challenging in the NLSY 79, in which the AFQT was administered in the summer and fall of 1980 to a cross-section of respondents aged 15 to 23. Those respondents with the most education at the time tended to perform relatively well on the test; they also tended to be the oldest.

In this paper, we present new estimates of the effect of schooling on AFQT performance for the youngest cohorts in the NLSY 79, who were between the ages of 15 and 19 when the test was administered. We address the issues raised above by comparing test-takers with birthdays near state-mandated cutoff dates for school entry. Because school entry laws specify an exact date by which entering first graders must reach age six (e.g., September 1), they generate sharp differences in average age at school entry—and thereby in average completed schooling—among students of nearly the same chronological age and, arguably, the same innate ability.[3] In principle, comparing test-takers with birthdays near school entry cutoff dates should reveal any causal impact of schooling on the AFQT.

In practice, small sample sizes in the NLSY 79 preclude us from making such "local" comparisons, however compelling. Instead, we instrument for schooling with "academic cohort"—the year in which a respondent should have entered first grade given her exact birth

---

[2] Indeed, even IQ was originally conceived as an age-graded measure of intelligence.
[3] Angrist and Krueger (1991) were the first to use school entry legislation to motivate an instrument for schooling.

date and state of birth—in a model that allows birth date to have a direct effect on AFQT performance. Since the AFQT was administered around the same time to all NLSY respondents, these controls for birth date are roughly equivalent to controls for age at the time of the test. Across-state variation in school entry cutoff dates then allows us to specify these birth date effects in a fairly unrestricted (non-monotonic) way. This is useful, given that season of birth—not just age alone—might be directly related to test performance.[4] In general, our model retains many of the desirable features of the idealized comparisons described above, but we pool across cohorts and states to achieve enough power to detect effects.[5]

We find evidence of an effect of schooling on the AFQT performance of minorities. For blacks, two-stage least squares (TSLS) estimates imply that an additional year of completed schooling by the test date raises AFQT performance by as much as 0.35 standard deviations—more than one-third of the black-white gap in AFQT scores. These findings are relatively convincing, since we detect no evidence for blacks that academic cohort is related to other determinants of AFQT performance, such as measures of family background at age 14. For native-born Hispanics, the results are more sensitive to specification, but qualitatively similar. However, we cannot detect an effect of schooling on the AFQT performance of whites in the NLSY with our identification strategy; we can neither rule out a large impact nor a lack of impact for this population.

Taken together, these findings are consistent with the achievement interpretation of the AFQT. Our estimates for minorities are in fact larger than those already documented in the literature for pooled samples (Neal and Johnson, 1996; Winship and Korenman, 1997; Hansen,

---

[4] This criticism was lodged by Bound, Jaeger, and Baker (1995) and Bound and Jaeger (2000) against the use of quarter of birth as an instrument for schooling by Angrist and Krueger (1991).
[5] In their study of maternal education and birth outcomes, McCrary and Royer (2003) have the sample sizes needed to apply a non-parametric estimation strategy in this context.

Heckman, and Mullen, 2004). Taken with this literature, this paper suggests that practitioners take care in interpreting studies in which the AFQT is correlated against economic outcomes, such as adult earnings. Such studies are more likely to estimate contribution of learned skill or achievement, as argued by Neal and Johnson (1996), than a contribution of fixed, innate ability. The large effect of schooling we find for minorities implies that studies that use AFQT to "explain" race differences in adult outcomes should be interpreted with particular caution, especially in light of the fact that many in the NLSY sample had not completed their education at the time of AFQT administration.

On a related note, our estimates suggest that schools play an important role in raising skill among minorities, even late in the school career. A long-standing but recently renewed debate concerns the relative importance of schools and families in explaining gaps in test performance between whites and minorities (Fryer and Levitt, 2004; Todd and Wolpin, 2004). We uncover an impact of schooling on AFQT scores for minorities in models that control for both family background at a given point in time and chronological age at the test date—one metric for the cumulative investments made in a child's human capital outside of school. Thus, even if low quality schools preclude test score gaps from closing as children progress through school (Fryer and Levitt, 2004), schools might nonetheless prevent these gaps from widening more dramatically as children age. While the present research does not attempt to assess the relative importance of schooling and family for achievement, the fact that our estimates show that age affects test performance separately from schooling – and the effects of both vary by race – indicates both could be important.

The paper proceeds as follows. In the next section, we give more background on the AFQT and previous attempts to identify its relationship to schooling in the NLSY 79. In Section

III, we motivate and describe our identification strategy, and in Section IV, we describe our

NLSY sample. Section V presents our findings, and Section VI gives several remarks on

interpretation. Section VII concludes.

II. Background

A. *NLSY Data on the AFQT*

The NLSY 79 is an ongoing panel survey of a nationally representative sample of

individuals born between 1957 and 1964. When first initiated in 1979, the NLSY had 12,686

participants between the ages of 14 and 22, more than 5,000 of whom constituted a supplemental

sample of Hispanics, blacks, and economically disadvantaged whites. In 1980, the U.S. military's

Armed Services Vocational Aptitude Battery (ASVAB) – a 10- section written exam administered

to all entering military recruits – was given to 11,914 of the NLSY respondents, around 94

percent of the original sample. NLSY participants were recruited to take the ASVAB through a

Department of Defense project that sought a nationally representative sample of youths from

which the Department could update the norms of the test from the previous World War II

reference population.

Each section of the ASVAB attempts to measure skills in a separate subject area. Some

subject areas are academic, and some are vocational.[6] Four of the 10 sections comprise the

AFQT, used by the military as the primary criterion to determine eligibility for enlistment and

"trainability" (Center for Human Resource Research, 2001, p. 94). Originally, AFQT percentile

scores were derived from a weighted sum of raw scores from two math and two verbal sections

of the test. Beginning in 1989, the Department of Defense changed which math sections were in

---

[6] The subjects are (1) general science; (2) arithmetic reasoning; (3) word knowledge; (4) paragraph comprehension; (5) numerical operations; (6) coding speed; (7) auto and shop information; (8) math knowledge; (9) mechanical comprehension; and (10) electronics information.

5

the AFQT, as well as the procedures for calculating the AFQT scores from the raw section scores.[7] In this paper, we use these "revised" AFQT scores, transformed into standard deviation units so that our results can be more readily compared to previous studies.

Several additional aspects of the AFQT and its administration in the NLSY are relevant for this analysis. First, the AFQT was given during the summer and fall of 1980, though the NLSY does not record the exact date that each respondent took the exam. While this time frame is narrow enough to precisely measure years of completed schooling at the test date, it is too wide to measure age at the test date with the same degree of precision.[8] In our analysis, we therefore control for date of birth, which can be precisely measured given the available data and should be highly correlated with age at the test date.

Second, the same test was administered to all NLSY respondents regardless of age or schooling attainment, though the test is designed to be appropriate only for individuals aged 17 and over (Center for Human Resource Research, 2001).[9] The AFQT is therefore not a school achievement test in the standard sense, as it is not intended to measure skills acquired from any given year of school. Thus, to the extent that the AFQT measures skills imparted through "nurture," it should be thought of as a cumulative measure of skill: an individual's score may be affected by human capital investments—both formal and informal—made from birth through the time of test administration.

---

[7] Originally, AFQT percentile scores were based on the sum of the word knowledge, paragraph comprehension, and arithmetic reasoning, plus one-half of the numerical operations raw scores. The post-1989 version of the AFQT first replaced numerical operations with math knowledge and changed the weighting of the math and verbal components in computing percentile ranks (NLS User Services, 1992, Tables C and D).

[8] We also unfortunately cannot address the problem of "summer setback"—the loss in skill that occurs when school is not in session. However, to the extent that summer setback affects children of all races similarly, and that the time of test administration is not related to race, comparisons of our results across race should remain meaningful. Fryer and Levitt (2004) provide evidence that summer setback is equally prevalent across race using a sample of young students.

[9] Though the youngest respondents in the NLSY took the AFQT, their scores were not used to create new norms for the test. The NLSY codebook suggests that considerable care be used in interpreting and using test scores for these respondents (Center for Human Resource Research, 2001).

Schooling and the AFQT: The Existing Evidence

This paper focuses on the role of later investments in raising AFQT performance. In particular, we examine whether an additional year of high school impacts skill as measured through the AFQT. We are interested in a model of the form

(1)		$T_i = \alpha + \theta\, S_i + f(V_i) + v_i$

where T_i is the AFQT score of individual i; S_i is completed schooling by the time of the test; $f(V_i)$ is some function of a vector of control variables, which might include measures of family background; and v_i is a mean independent error term. Most researchers have also included crude controls for age, such as indicators for calendar year birth cohort, in the vector of controls. The parameter θ gives the marginal effect of schooling on AFQT performance.[10]

In practice, it is impossible to estimate model (1): $f(\cdot)$ is an unknown function, and the entire vector of appropriate control variables may be unobservable to the researcher. Instead, the existing literature has focused on a restricted version of the model:

(2)		$T_i = \alpha + \theta\, S_i + \beta' X_i + \varepsilon_i$,

where $X_i \subset V_i$ is some subset of controls. A new error term, ε_i, now includes unobservable determinants of test performance (such as fixed, innate ability) that are potentially correlated with S_i. Thus, the parameter of interest, θ, will not necessarily be identified through simple least squares regression.

Researchers have addressed this identification problem in several ways. One approach, employed by Herrnstein and Murray (1994) and Winship and Korenman (1997), assumes that

---

[10] This is a restricted version of a more general model discussed in Hansen, Heckman, and Mullen (2004). We restrict ourselves to this model for several reasons. First, this specification has formed the basis of most other studies of the relationship between AFQT performance and schooling, as discussed here. Second, our estimation strategy is sufficient to identify one parameter; more general models would not be identified.

schooling is randomly assigned conditional on a "pretest," or a measure of IQ at a younger age.

An alternative approach is to find an instrument for schooling at the test date, then estimate

model (2) using two-stage least squares.   For example, Neal and Johnson (1996) use quarter of

birth as an instrument in estimating the effect of schooling on the AFQT. They argue that

quarter of birth is related to schooling through state school entry legislation: if a child's birthday

is after the cutoff date for school entry (e.g., September 1), he must wait an additional academic

year before starting first grade, thereby falling a full grade behind slightly older peers born during

the same calendar year. This within-cohort grade difference will be present when individuals are

of school age and will persist if dropout decisions are responsive to compulsory schooling

regulations (Angrist and Krueger, 1991).[11]

Both of these reduced-form approaches have generated similar estimates of effect of

schooling on AFQT performance. Using the sub-sample of NLSY respondents for whom early

IQ is recorded, Winship and Korenman (1997) estimate that an additional year of schooling

raises AFQT performance by 4 percentile points, or approximately 0.1 standard deviations.

When they take into account measurement error in early IQ and education, they estimate that the

effect could be larger—on the order of 5 to 6 percentile points, or 0.15 standard deviations.[12]

Applying the instrumental variables strategy discussed above to the youngest NLSY cohorts, the

majority of whom were still enrolled in school when the AFQT was administered, Neal and

Johnson (1996) estimate that a year of schooling is associated with a gain of 0.22 to 0.25 standard

deviations in AFQT scores.

---

[11] Individuals with birthdays immediately after the school entry cutoff date reach the minimum age of school exit having completed less schooling than individuals born right before.

[12] A similar approach is taken by Herrnstein and Murray (1994), as discussed above. However, in correcting several errors in Herrnstein and Murray's analysis (e.g., failing to control appropriately for age and problems with coding of missing data), Winship and Korenman (1997) find considerably larger effects of schooling on the AFQT.

III. Identification Strategy

A. Motivations

The empirical strategy undertaken in this paper has similar motivations to Neal and Johnson (1996). However, we attempt to identify the effect of schooling on the AFQT using only that component of season-of-birth variation in schooling that is in fact driven by school entry legislation. As shown in Table 1, school entry cutoff dates faced by the youngest NLSY cohorts were concentrated in the third and fourth quarters of the calendar year.[13] It therefore seems inappropriate to base comparisons on children born earlier in the year, particularly if season of birth is not randomly assigned.

To demonstrate more plainly the effects of school entry legislation, Figure 1 plots average highest grade completed at the test date by date of birth for native-born NLSY respondents born between 1961 and 1964.[14] The panels correspond to the modal cutoff dates for school entry, shown in Table 1: September 1 (Panel A), September 30/October 1 (Panel B), December 1 (Panel C), and December 31/January 1 (Panel D).[15] Respondents with birthdays between any two adjacent vertical lines can be thought of as belonging to the same "academic cohort": if entry legislation were binding and grade progression perfect, they should have completed the same number of grades at a given point during the school career. Using this rule, the first academic cohort depicted should have entered school in 1967 and should have therefore completed 12th grade plus one year of postsecondary education (13 grades) by the summer of 1980, when the AFQT was administered; we refer to this cohort as the 1967 academic cohort

---

[13] These cutoff dates were acquired by the authors from legal archives. See Appendix A. Dates are adjusted for consistent interpretation: a child residing in a given state is able to begin first grade if she has reached the required age no later than the date given.

[14] Lacking information in the NLSY on age at school entry, we cannot measure number of years spent in school.

[15] Respondents have been assigned school entry dates on the basis of state of birth, since state of school entry is not recorded in the NLSY. In the few instances where state of birth is missing, we use state of residence to assign cutoff dates to respondents, provided that the respondent reports having never moved.

below. Similarly, we refer to the next four academic cohorts as the 1968 through 1971 academic cohorts, respectively.

The figure shows that compliance with school entry legislation is far from complete in this sample. A substantial fraction of respondents are "below grade," having completed fewer grades than would be predicted given their state and date of birth. Respondents might be below grade for three reasons. The first is delayed entry into school: school entry legislation pertains to the earliest age at which a child *may* enter school; in states with minimum compulsory schooling ages of greater than six years, parents may opt to delay the school start of eligible children. Second, a child may start school when eligible but may repeat a grade later. Finally, if state compulsory schooling legislation permits, some respondents may have dropped out of school, failing to progress to the next grade as predicted.[16]

As predicted, however, the figure shows that the largest differences in completed schooling occur among individuals with birthdays near the school entry cutoff dates in their states of birth. These differences are most evident for the younger cohorts depicted in the graph. There is also no apparent discontinuity of completed schooling around December 1 or January 1 for the sample of individuals born in states with September 1 cutoff dates for school entry, and *vice versa*. This is what we would expect: if subject to a September 1 regulation, a child would be able to begin school if born on August 31 or September 1 but would not be able to begin school if born on September 2 or later during the same calendar year. In a state with a December 1 cutoff, these abrupt changes should occur among individuals with birthdays at the beginning of December, not those born earlier in the academic year.

We can formalize this intuition by estimating the differences-in-differences (DD) model

---

[16] Unfortunately, we cannot examine these alternative explanations with the variables available in the NLSY.

(3) $S_i = \delta_0 + \vartheta_1 \cdot POST_i + \delta_1 \, CUTOFF_i + \mu \, POST_i \cdot CUTOFF_i + \omega_i$

where $POST_i$ is an indicator equal to one if individual i is born in the month after some specified

date (e.g., September 1); $CUTOFF_i$ is an indicator equal to one if this specified date was the cutoff

date for school entry faced by individual i; and $POST_i \cdot CUTOFF_i$ is the interaction of these

variables. If school entry legislation were the only force behind trends in completed schooling,

then it should be the case that $\mu = -1$ and $\vartheta_1 = 0$: having a birthday after the specified date only

reduces completed schooling if that date is an entry cutoff.

A special case of this model arises when we focus only on the subsample of respondents

for which $CUTOFF_i = 1$:

(3') $S_i = (\delta_0 + \delta_1) + (\vartheta_1 + \mu) \cdot POST_i + \omega_i$

We refer to this model as a regression discontinuity (RD) specification. Equation (3') is a simple

"pre-post" comparison of individuals with birthdays near school entry cutoff dates, analogous to

the stylized comparisons described in the introduction.

Estimation of (3) may be preferable. There have historically been seasonal cycles in the

total number of births in the United States, with the peak in total births in September and the

trough around April or May. While the exact cause of these seasonal fluctuations in fertility is

unknown, the fertility of lower socioeconomic status groups has historically been more volatile

(Lam and Miron, 1991; Kestenbaum, 1987; Seiver, 1985; Warren and Tyler, 1979). Month of

birth (i.e., $POST_i$) might therefore proxy for otherwise unobservable dimensions of family

background, which might impact grade retention, delayed school entry, or dropout (in which case

$\vartheta_1 \neq 0$).

11

Table 2 gives estimates of equations (3) and (3') for four different sets of school entry cutoff dates (those depicted in Figure 1) and for blacks in our estimation sample, which spans the 1968 to 1970 academic cohorts (see below).[17] Column (1) gives our estimates of the simple RD specification, equation (3'). As expected, individuals with birthdays in the month after the cutoff date have significantly lower levels of schooling as teenagers than those born in the month prior, regardless of the cutoff date. Estimates of the effect of being born after the cutoff date on schooling range between 0.619 to 1.445 fewer years of schooling completed by the summer of 1980. Column (3) shows that these local differences in completed schooling remain in the DD specification, which allows for a direct effect of month of birth on schooling. Here, estimates of μ range between -0.48 and -1.198 and are statistically different from zero in three out of four cases. In no case can we reject that $\vartheta_1 = 0$, though this could primarily be a reflection of small sample sizes.

B.   *Illustration*

These abrupt changes in average highest grade completed with birthday provide the basis of our strategy for identifying the relationship between schooling and the AFQT.[18] *A priori*, it seems believable that non-random sorting around the cutoff date would be uncommon, in which case individuals born in adjacent months around the cutoff will be on average of the same ability, even though they would have completed significantly different levels of schooling. This motivates a first-stage model such as (3'), where $POST_i$ is used as an instrument for schooling in a simple model of the schooling-test score relationship, such as that given in (2).

---

[17] In all regressions, we control for a "comparison" fixed effect. In Panels A to C, this is equivalent to a year of birth fixed effect. In Panel D, we include an indicator for whether an individual was born in December 1963 or January 1964.

[18] The regression-discontinuity approach is employed by Cahan and Cohon (1989) to separately identify the effects of schooling and age on achievement test scores in Israel.

However, just as was the case with schooling, there are several reasons why $POST_i$ might be directly related to test scores. As described above, month of birth might affect test scores through its correlation with otherwise unobservable dimensions of family background. Further, birth month will be highly correlated with age at the test date. Small changes in age might have a direct effect on test performance through either biological "maturation" or the cumulative effects of investments in human capital outside of school.

These observations motivate a richer model of the test score-schooling relationship, where month of birth can exert a direct effect on test scores:

(2') $T_i = \alpha_0 + \theta \, S_i + \psi_1 \cdot POST_i + \alpha_1 \cdot CUTOFF_i + \varepsilon_i$

Here, the DD model in equation (3) serves as a first stage. The effect of schooling on the AFQT will be identified provided that the *interaction* of birth state and birth month ($POST_i \cdot CUTOFF_i$) is unrelated to innate ability or unobserved determinants of AFQT performance more generally. Intuitively, the exclusion restriction will hold if families do not selectively "choose" birth timing on the basis of the placement of school entry cutoff dates. Notice that this model can be identified given the variation in cutoff dates across states.

The remaining columns of Table 2 illustrate this DD identification strategy and its more naïve RD counterpart. We first give the reduced-form effects of school entry legislation on test scores, then the corresponding TSLS effect of schooling on AFQT performance. In the RD comparisons (columns (1) and (2)), TSLS estimates of the effect of an additional year of schooling on AFQT scores range between 0.027 and 0.319 standard deviations. The DD estimates are in general smaller, ranging between -0.139 and 0.309, and less precisely estimated. While most of the reduced-form and TSLS estimates are of the expected sign and a reasonable magnitude, in no case do we have enough power to detect an effect of schooling on the AFQT.

Estimates of model (2′) also provide no evidence of a direct effect of birth date on test performance, but this too might be an artifact of small sample sizes.

Table 2 thus shows that there is limited scope to make local comparisons in this application. The remainder of this section describes an alternative approach to exploiting the policy-induced discontinuities of schooling in birth date to identify the relationship between schooling and the AFQT. Our model retains the desirable properties of the comparisons made above but pools across cutoff dates and cohorts to achieve enough power to detect effects.

C.   Empirical Framework

The model of interest is an extension of (2′):

(4) $T_{is} = \theta\, S_{is} + \psi(D_{is}) + \beta' X_{is} + \alpha_s + \varepsilon_{is}$

where D_{is} is the birth date of individual *i* from state *s*; α_s is a state fixed effect; and all other variables are as previously defined. As discussed above, an individual's age at the test date will be highly correlated with D_{is}.[19] In our estimates, we allow for $\psi(\,\cdot\,)$ to be either a smooth function (e.g., a polynomial) or an unrestricted function (e.g., a linear function of interaction of year and month of birth). While the former is more intuitive—age or maturation itself will not have abrupt or discontinuous impacts on test performance—the latter is preferable, since birth date might also correlate with family background.

As described above, schooling at the test date might still be related to unobservable determinants of AFQT performance, even conditional on age and family background. We therefore propose to use academic cohort as an instrument for S_{is}. Formally, we assign respondents to academic cohorts using the following rule:

---

[19] We therefore at times refer to $\psi(\,\cdot\,)$ as an "age control" in the following discussion.

$$(5) \qquad k = \begin{cases} y_{is} + 5 + 1[d_{is}^* > 32] & if \ d_{is} \le d_{is}^* \\ y_{is} + 6 + 1[d_{is}^* > 32] & if \ d_{is} > d_{is}^* \end{cases}$$

where y_{is} represents year of birth and $d_{is} = 1,\dots,365$ represents day of birth within the calendar year (which together define an individual's date of birth, D_{is}), and d_{is}^* stands for the day of the year on which the school entry cutoff date fell in state s during the calendar year when individual i turned age six.[20] The term $1[d_{is}^* > 32]$ is an indicator equal to one if the relevant cutoff date is after February 1 during the calendar year. To take an example, equation (5) assigns an individual born in 1962 in a state with a September 1 cutoff date to the 1968 academic cohort if born on or before September 1, and to the 1969 academic cohort if born after September 1.[21]

One first stage model is just a generalization of (3):

$$(6) \qquad S_{is} = \sum_{k=m}^{M} \mu_k' C_{is}^k + \vartheta(D_{is}) + \gamma' X_{is} + \delta_s + \omega_{is} \,,$$

where $C_{is}^k; k = m,\dots,M$ are a series of indicators for academic cohort, δ_s is a state fixed effect, and all other variables are as previously defined. Coefficients on the academic cohort indicators, μ_k, give the average impact on schooling of being assigned to enter first grade in year k, holding constant family background, fixed differences in completed schooling across states, and what would have otherwise been expected for individuals born at the same time. Notice that C_{is}^k is simply a generalization of the interaction term $(POST_i \cdot CUTOFF_i)$ employed as an instrument in the DD specifications in Table 2. However, here we extend the definition to individuals born

---

[20] We do take leap years (e.g., 1964) into account when defining this variable. For leap years, $d_{is} = 1,\dots,366$, and d_{is}^* is redefined as $d_{is}^* + 1$, where appropriate. With a few exceptions (e.g., North Carolina and Delaware), the school entry dates faced by all respondents born in the same state were identical, so $d_{is}^* = d_s^*$ for all individuals i in state s.

[21] The term $1[d_{is}^* > 32]$ is needed because some states have cutoff dates in the first quarter, as shown in Table 1. For example, consider a respondent born on January 1, 1963 in a January 1 state. He will turn six years of age on January 1, 1969, and should therefore legally be able to begin first grade in the fall of 1968. Without addition of the term $1[d_{is}^* > 32]$, he would be assigned to the 1969 academic cohort.

further away from the cutoff date. Assuming that the dummy for the oldest academic cohort in

the sample (m) is omitted for the purposes of identification, perfect compliance with the law

predicts that $\mu_{m+j} = -j$, or that every one-year increase in academic cohort lowers completed

schooling at the test date by on average one year.

More generally, if the marginal effect of academic cohort is constant across all cohorts

under consideration, it will be the case that $\mu_{m+j} = \mu \cdot j$, where $\mu < 0$. Imposing this restriction

on equation (5), we arrive at an alternative model for schooling:

$$(6') \qquad S_{is} = \mu\, C_{is} + \vartheta(D_{is}) + \gamma'\, X_{is} + \mu_s + \omega_{is},$$

where $C_{is} = \sum_j j \cdot C_{is}^{m+j}$. We focus on this alternative model in the analysis below, and the

restrictions are, in practice, only rarely rejected. The parameter μ is interpreted like it was in

equation (3): μ gives the reduced-form effect of being born after the school entry cutoff,

controlling for date of birth.

We use equations (6) and (6') as first-stage models in estimating equation (4) with TSLS.

When $\vartheta(\cdot)$ is specified as a smooth function, these equations reduce to parametric RD

models.[22] When $\vartheta(\cdot)$ consists of a series of dummy variables for birth date, these equations are

the analogs to the DD models presented above. It is once again the across-state variation in

cutoff dates that allows us to identify an unrestricted effect of birth date while simultaneously

identifying an effect of schooling on test scores.

In general, academic cohort will be a valid instrument if it exhibits no relationship

with ε_{is}, conditional on the age controls and other observable characteristics. The intuition here is

---

[22] As in any parametric RD framework, it is useful to control for a smooth function in the index determining selection into the treatment (i.e., birthday) if not making a comparison local to the selection threshold; failure to do so may yield spurious discontinuities in schooling and AFQT performance (Cook and Campbell, 1979).

the same as in the illustration: the identification assumption requires that there is no sorting around school entry cutoff dates, i.e., no "gaming" of school entry legislation through birth timing. It is important to note here that delayed school entry or grade repetition does not represent a violation of this identification assumption; it is precisely this type of variation in schooling that suggests the need for an instrument. Rather, our identification strategy requires that only time of birth be exogenous within a narrow window of the cutoff date. While this assumption cannot be directly tested, below we test for a relationship between academic cohort and observed correlates of test scores, such as parental education. We also perform standard tests of over-identifying restrictions when TSLS estimates are based on (6).

IV. Data

A. Sample

Because it is the subpopulation for which the relationship between school entry laws and schooling is both theoretically and empirically the strongest, we restrict attention to the youngest NLSY respondents in the analysis that follows. Specifically, we focus on individuals who, given their birth date and state of birth, should have entered first grade between 1968 and 1970.[23] These respondents would have been between the ages of 15 and 19 during the summer of 1980 and, therefore, would have been relatively more likely to have been enrolled in school during the prior academic year. These academic cohorts correspond to the three fully observed academic cohorts in Figure 1, which should have completed grades 10 through 12 by the test date if on grade.

---

[23] Because the full 1971 academic cohort cannot be observed, as suggested by Figure 1, we omit individuals in this cohort from our analysis. In practice, we drop only individuals born at the end of 1964 as a result of this selection rule. It might in fact be appropriate to omit these individuals from the analysis, given evidence of "non-random self-selection at the edge of the age range" for the survey (Center for Human Resource Research, 2001, p. 17).

Otherwise, we limit our sample to respondents about whom information on day of birth is not missing, and who were born in states in which school entry cutoff dates are set by the state, not local discretion (see Table 1). So that we have sample sizes sufficient to conduct the analysis separately by race, we use data on blacks and Hispanics from the supplemental NLSY sample. The resulting sample consists of 3300 observations and is described in detail in Appendix B.

B.   *Summary Statistics*

Table 3 presents the summary statistics of key variables for our sample, both overall and stratified by race. The table first presents means of schooling at the test date, age, and several measures of AFQT performance (Panel A). These measures include the revised (1989) AFQT percentile score, the standardized version of this score (used in the regressions below), and the standardized scores on the four ASVAB components that comprise the AFQT. Overall, the sample was on average slightly over age 17 on July 1, 1980, and had on average completed 10.64 years of schooling when the test was administered. The average sample respondent also scored in the 44th percentile on the test, or about a one-quarter of a standard deviation below average for the population as a whole.[24]

The remainder of Panel A shows differences across race in average AFQT performance and schooling at the test date. Blacks and Hispanics have completed less schooling at the time of the test, despite being the same age as white respondents, on average. This lower level of educational attainment might arise from differences across race in the incidence of grade

---

[24] This below-average performance is due to the fact that test norms are based on the 1957 through 1962 birth cohorts (Center for Human Resource Research, 2001), but our sample contains the youngest respondents in the NLSY.

retention or delayed school entry or in dropout behavior.[25] Our sample also reproduces the well-known gaps in test scores across race. White respondents score on average in the 49th percentile, while black and Hispanic respondents have AFQT scores in the 20th and 30th percentiles, respectively. These percentile differences translate into a 1.1 standard deviation test score gap between blacks and whites and a 0.7 standard deviation test score gap between Hispanics and whites. The test score gaps are relatively narrow for mathematical components of the ASVAB (arithmetic reasoning and math knowledge) and relatively wide for the verbal components of the test battery (word knowledge and paragraph comprehension).

Alongside the sample means for each variable are F-statistics that correspond to a test for significant differences in means across academic cohorts. Underlying models control for state of birth and year of birth by month of birth fixed effects, the preferred DD specification. In Panel A, these F-statistics test the strength of the instruments in the preferred reduced-form regressions. In Panel B, they provide a crude test of the identifying assumption of the model: as described above, if academic cohort were a valid instrument, we might expect it to have little predictive power with regard to other observed correlates of test performance, such as family background. Panel C suggests that there is indeed enough variation in cutoff dates across states to estimate the DD specification, though most sample respondents were born in states with fourth quarter cutoffs.

Focusing first on Panel A, we find remarkable heterogeneity across race in the strength with which academic cohort predicts schooling and test performance. Academic cohort is strongly related to schooling in 1980 in the sample overall (F-statistic=10.7), even controlling for year of birth by month of birth fixed effects. The relationship is strongest for blacks (F-

---

[25] Grade retention is likely to play the most important role here, though this cannot be confirmed with the available NLSY data. At least in more recent cohorts, delayed school entry appears to be more common among whites and grade retention more common among minorities (National Center for Education Statistics, 2000).

statistic=22.2) and weakest for native-born Hispanics (F-statistic=2.7). Academic cohort also has explanatory power only with regard to the AFQT performance of blacks. Given this racial heterogeneity and the history of using the AFQT to understand race-specific differences in outcomes, we conduct the entire analysis below separately by race.

It is rare that we find a significant correlation between academic cohort and basic measures of family background (all measured at age 14), as shown in Panel B. For the black sub-sample, all F-statistics on academic cohort are around 1 or below. For whites, however, there is a significant relationship between the instrument and maternal education (F-statistic=5.0). Even after controlling for maternal education in our regressions, we might therefore be concerned that the instrument is correlated with unobservable determinants of AFQT performance and is therefore not valid. Similarly, academic cohort is strongly related to the likelihood that a native-born Hispanic respondent has migrated from his state of birth. To the extent that migration takes place before school entry and involves states with different cutoff dates, this migration could generate misclassification in academic cohort and obscure the first-stage relationship. This migration might, however, be a selective response to the cutoff dates.

Are our samples of whites and Hispanics just unrepresentative draws from the population, or do these tests evidence true "gaming" of the cutoff dates? If the former, we would merely interpret our findings for whites and Hispanics with due caution; if the latter, we might be concerned about the credibility of the comparisons made for blacks as well.

To investigate these possibilities, we performed a similar exercise to that done in Table 3 using data on the NLSY cohorts from the 1970 and 1980 decennial censuses.[26] Specifically, we retained data on all whites, blacks, and Hispanics who would have entered school between 1968

---

[26] We draw our samples from the 1970 Form 2 (15 percent) state file and the 1980 state (A) file. The 1970 sample represents 1 percent of the population, and the 1980 sample represents 5 percent of the population.

and 1970 (as in our NLSY sample) and who were born in states with September 30, October 1, December 31, or January 1 cutoff dates (as in Table 1). We thus limit ourselves to a group of states for which we are able to define academic cohort fairly precisely using information on quarter of birth and age on the census date. In 1970, we furthermore matched children with their mothers to generate two measures of family background reported in Table 3 – maternal education and family size (mother's number of children ever born) – though at a younger age than they are reported in the NLSY.[27]

Table 4 gives the results of this exercise. The F-statistics correspond to a test for significant differences in means across academic cohorts, now adjusted for state of birth and year of birth by quarter of birth fixed effects. We see in the census the same pattern of results for schooling that we observed in the NLSY: academic cohort is significantly related to highest grade attended in 1970 (Panel A) and to both highest grade attended and highest grade completed in 1980 (Panel B), but only for non-Hispanic whites and blacks. As was the case in the NLSY, there is some evidence for Hispanics of a relationship between academic cohort and the likelihood that a respondent has migrated from his state of birth. This relationship is statistically significant only in the 1970 sample.[28] In the census, we also no longer find a significant relationship between academic cohort and maternal education for whites that we observed in the NLSY, though we do find a marginally significant relationship between academic cohort and maternal fertility.[29]

---

[27] The match rate is much smaller in the 1980 census and the matches much more selective, given that the relevant cohorts were between the ages of 15 and 19.

[28] The census results for Hispanics in 1970 (where migration is measured as of ages 5 to 9) for Hispanics in the NLSY (where migration is measured as of age 14) are in fact fairly similar in magnitude. In the 1970 Census, the coefficient (standard error) on academic cohort in the migration regression for Hispanics is -0.127 (0.057). In the NLSY, the analogous coefficient is -0.185 (0.043).

[29] In the 1970 census, the coefficient (standard error) on academic cohort in the maternal education regression for whites is -0.048 (0.045). In the NLSY, the analogous coefficient (standard error) was -0.570 (0.254). The coefficient (standard error) on academic cohort in the fertility regression for whites is 0.102 (0.058).

We draw two conclusions from this exercise. First, the NLSY sample appears to be a rare, unrepresentative draw from the underlying population of native-born whites. This conclusion is bolstered by the fact that we continue to find no evidence of a relationship between academic cohort and other background characteristics for blacks and Hispanics. Although we present results for whites in the analysis below, we discount our findings for this subpopulation; in our NLSY sample, school entry legislation essentially does not provide an experiment with which we can credibly identify the effect of schooling on the AFQT.

Second, the migration of Hispanic children appears to occur before they would have entered school. Below, we address this problem by presenting an alternative set of regression results for non-movers and where academic cohort is assigned using state of residence. We interpret these findings with considerable caution, since we cannot rule out that migration was a selective response to school entry laws.

V.   Results

Table 5 gives the reduced-form estimates from the just-identified model (equation (6') and its analog for AFQT scores), separately by race.[30] The specifications differ in which controls for birth date are employed and whether controls for family background (at age 14) have been included.[31] Table 6 gives the corresponding TSLS estimates from this just-identified model, along with OLS estimates for the purposes of comparison.

---

[30] In all reduced-form and subsequent regressions, we present standard errors robust for clustering within state.
[31] Background controls include mother's highest grade completed, father's highest grade completed, number of siblings at age 14, an indicator for whether the respondent has migrated from his state of birth, and an indicator for whether the respondent is female. Where family background variables are missing, they have been imputed with sub-sample specific means (i.e., separate means are calculated for cross-sectional sample of black females, the cross sectional sample of black males, the supplemental sample of black males, etc.). We also control for indicators for whether background variables have been imputed.

We begin with a simple specification, which includes no controls for age or background. As shown in column (1) of Table 5, the effect of school entry laws is large, ranging in value between 0.76 and 0.92 fewer years of schooling completed for every additional academic cohort, depending on race. Black and white respondents who are predicted to have entered school later also tend to have lower test scores. For these subsamples, TSLS estimates of the effect of an additional year of schooling on the AFQT range between 0.12 and 0.13—effect sizes that are comparable to those found in past studies. We present these estimates only for completeness, since academic cohort is strongly correlated with birth date, which has been omitted from the model even though it might be directly related to AFQT performance.

The remaining columns in Tables 5 and 6 add controls for different specifications of date of birth to the model and alternatively exclude and include family background controls. We first show an RD specification, where we control for birth date (or age as of July 1, 1980) with a fourth-order polynomial (column (2)).[32] This simple, smooth control for age has a dramatic impact on both of the reduced-form models. In the schooling models for white and Hispanic respondents, these terms reduce the coefficient on academic cohort by around 50 percent, though the instrument remains a strong predictor of highest grade completed in both cases. Consistent with the earlier illustration, controlling for birth date has less of an impact on the reduced-form schooling model for blacks. However, regardless of race, controlling for a smooth function in age eliminates the statistical relationship between the AFQT and academic cohort. The coefficients on academic cohort fall in magnitude and cannot be distinguished from zero. With controls for family background (column (3)) the reduced-form coefficients are slightly larger for blacks and Hispanics but are still not distinguished from zero.

---

[32] The results are very similar when we control for only a linear term in birth date. They are also similar to results from models where we include the interactions of birth state indicators with the linear age term.

Age effects, the coefficients of which are not shown in Tables 5 and 6, turn out to be highly significant for whites and Hispanics. Conditional on academic cohort, completed schooling rises with age, though more steeply for whites and Hispanics than for blacks. Average test performance also rises with age, though only for whites. One interpretation of these age effects is that more informal human capital investment occurs for whites, having a strong effect on the likelihood that whites progress through school normally and perform well on tests. Another is that age relative to one's classmates has a direct effect on achievement, a phenomenon that has been suggested by other research (Datar, forthcoming).[33] Regardless of what generates these positive age effects, however, it is important to note that age effects for test scores are negative in the OLS specifications: Individuals who are old in their grade are negatively selected. The fact that more plausible positive age effects are found in TSLS estimates is consistent with the view that our approach identifies exogenous variation in schooling.[34]

The model presented in columns (2) and (3) is restrictive in light of the many ways in which date of birth might impact both schooling and AFQT performance, outlined above. In particular, we may have omitted family background attributes correlated with season of birth. The specification in column (4) adds a series of dummy variables for month of birth to capture "season of birth" effects common to all cohorts under observation. Notice that this specification allows AFQT scores to be directly affected by variables that have previously been used as instruments for schooling (Neal and Johnson, 1996). Even conditional on month effects, academic cohort has a significant effect on the schooling and a marginally significant impact on the AFQT performance of blacks. This relationship is strengthened with controls for family background (column (5)).

---

[33] We return to a discussion of relative age, and its potential effects on our estimates, in the next section.
[34] The age effects are not reported to conserve on space. Coefficient estimates are available from the authors upon request.

In the final two columns we present the least restrictive age controls, where we include the interaction of month of birth and year of birth fixed effects. This is the preferred specification and can be thought of as a generalization of the DD models presented in Table 2. For blacks, the reduced-form DD and corresponding TSLS estimates closely resemble those given in Table 2 but are more precisely estimated. The first-stage relationship between highest grade completed and academic cohort is strong, with an F-statistics on academic cohort of 22.2. The reduced-form relationship between academic cohort and AFQT scores is also marginally significant, with an F-statistic of around 3.3. Family background controls (column (7)) also have very little impact on the coefficient estimates. The corresponding TSLS estimates (columns (6) and (7) of Table 6) show that a one-year increase in completed schooling at the test date is associated with a 0.3 to 0.35 standard deviation increase in AFQT performance.

For the white and Hispanic subsamples, the results with more restrictive age controls are less robust. For whites, controlling for birth date fixed effects diminishes the coefficient on academic cohort in the first stage. Adding family background makes the first-stage relationship insignificant (column (5) of Table 5). As anticipated, the corresponding reduced-form and TSLS estimates for test scores are also very imprecise. In specifications with family background controls, the TSLS coefficient on schooling is negative but so imprecisely estimated that we cannot rule out large positive effects.[35] We place little emphasis on these results, since, as we demonstrated above, school entry laws essentially do not provide an experiment with which we can identify the effect of schooling in the NLSY.

---

[35] One might be concerned that negative point estimates for whites were coming from "ceiling effects." In particular, although an additional year of schooling may raise the *latent* skill of all individuals sampled, this gain in skill will not be revealed to the extent that an individual would have otherwise scored close to (or at) the maximum. About 1.7 percent of the white sample scored in the 99th percentile on the AFQT, and 13 percent scored at the maximum on at least one of the four ASVAB tests. Reduced-form models estimated with Tobit look very similar to their OLS counterparts. (These estimates are available from the authors upon request.)

For Hispanics, the reduced-form regressions with unrestricted age effects are also sensitive to the inclusion of family background controls. However, in this case, controlling for family background enables us to bound the first-stage relationship between academic cohort and schooling away from zero (F-statistic=4.2). We also find some evidence of a relationship between academic cohort and AFQT performance (F-stat=2.3). When these reduced-form estimates are combined through TSLS, we estimate that a one-year increase in schooling at the test date raises AFQT scores by around 0.34 standard deviations (column (4) of Table 6, Panel C). The effect size is even larger in the sample of non-movers and when the models are re-estimated assigning cutoff dates using state of residence at age 14. As shown in Table 7, the effect size is around 0.57 for Hispanics in these alternative specifications and highly significant. However, we are hesitant to draw particularly strong conclusions from these estimates, since they appear to be very sensitive to controls for both migration and family background.

Returning to the original model, we find that TSLS estimates from over-identified specifications (Table 6), where dummies for academic cohort are used as instruments, are fairly similar to those from just-identified specifications.[36] For blacks, the TSLS estimates from the over-identified model are around 0.05 standard deviations higher than in the just-identified model and on the margin of statistical significance in the preferred specification (columns (6) and (7)). For Hispanics, the estimates from the over-identified model are almost identical to those in the just-identified case. With the exception of the naïve specification in column (1), in no case can we reject the over-identifying restrictions of the model at conventional significance levels.[37]

---

[36] Although we rarely reject the restrictions equation (6') places on equation (6). F-statistics on the model restriction are suppressed due to space constraints but are available from the authors upon request.

[37] The test statistic is the number of observations, multiplied by the R-squared from a regression of the TSLS residuals on the instruments and other observable characteristics. With two instruments (and one over-identifying restriction), the test statistics are drawn from a χ^2 distribution with one degree of freedom.

Collectively, these results suggest that the AFQT measures school achievement, not some notion of fixed, innate ability. For blacks and Hispanics—the two subpopulations for whom school entry legislation appears to provide credible identifying variation in schooling—we find that an additional year of schooling raises AFQT scores by up to 0.35 standard deviations. When the minority sub-samples are pooled, estimates of the schooling effect are of a similar magnitude but more precise.[38] We unfortunately cannot say anything conclusive about the role of completed schooling in the AFQT scores of whites. In the preferred DD specification, we cannot rule out a similar schooling effect; however, the estimates are so imprecise that we also cannot rule out negative effects of the same magnitude. Given that the NLSY appears to be a "bad draw" for the purposes of implementing our identification strategy, we view this as a call for further research in this question, rather than an indictment of the strategy itself.

VI.   Discussion

In this application, age has represented not only developmental maturation but also the cumulative effects of human capital investments made outside of school. "Aging" and schooling therefore present competing explanations for the evolution of test scores as children grow older. Not surprisingly, separately identifying the effects of schooling and age on test performance has long been a challenge for researchers (Ceci, 1991). Attempts to do so are important, however, for understanding the role that schools play in building skill. More to the point of this paper, acquiring a credible estimate of the contribution of schooling to performance on the AFQT in

---

[38] In the just-identified model with family background controls, state of birth fixed effects, and year of birth by month of birth interactions, we estimate that a one-year increase in completed schooling at the test date will raise AFQT performance by 0.379 standard deviations (clustering robust standard error=0.196) in the minority sub-sample (n=1575). In the comparable just-identified model including the quartic in age and month of birth fixed effects, we estimate a schooling effect of 0.327 (clustering robust standard error=0.146). Similar results are obtained in the corresponding over-identified models. (Standard errors are clustered on state of birth.)

the NLSY 79 has implications for the interpretation of countless studies where this test has been used as a proxy for skill.

A key difference between this study and previous investigations of the relationship between schooling and the AFQT, and test scores more generally, is the relatively unrestricted way in which we have allowed age at the test date to enter the model. Forcing these effects to be discrete, infrequent, and arbitrarily placed—e.g., with year of birth fixed effects, as common in the literature—is highly unintuitive. Unlike the previous literature, we have therefore allowed small changes in age to have a direct effect on AFQT performance. Second, we have allowed these age effects to differ by race. This too is intuitive, since race has historically been highly correlated with the resources a family has available to devote to investments in a child's human capital. We expected age gradients in test scores to be stronger for whites, and this is precisely what we found.

Finally, with variation in school entry cutoff dates across states, we were able to control for age in a fairly unrestricted fashion. In our least restrictive specifications, we allowed for a direct effect of season of birth on schooling and test scores. Although the non-monotonic (and discontinuous) nature of these controls conflicts with the intuition described above, age at the test date is highly correlated with birth date in the NLSY 79. If birth date is related to unobserved correlates of test performance, forcing age effects to be monotonic or overly smooth might yield misleading inferences about the role of schooling. Interestingly, in specifications where we included month of birth fixed effects, either alone or interacted with birth year, we still found an impact of school entry legislation on highest grade completed. For blacks, and to a lesser extent for Hispanics, we also uncover a corresponding relationship between academic cohort and AFQT scores in these specifications.

On that note, it is worthwhile and interesting to compare our results to those based on the standard quarter-of-birth instrument for schooling, which is ostensibly based on the same source of policy variation. Table 8 performs this exercise. The estimates based on the quarter of birth instrument differ notably from those presented in this paper. Consider first the model in (3). Though Neal and Johnson (1996) do not present separate results by race, their results correspond most closely to this specification, which does not control for family background or state of birth, and includes only year of birth fixed effects as a crude control for age.[39] In this specification, an additional year of schooling is estimated to increase AFQT performance by 0.18 standard deviations for whites, 0.38 standard deviations for blacks, and 0.58 standard deviations for Hispanics. All estimates can be statistically distinguished from zero.

The differences in results across specification can be traced, at least in part, to a strong direct effect of season of birth on test scores. The data strongly reject the exclusion restriction in the quarter-of-birth models, as evidenced by F-statistics on the year of birth by month of birth interactions in the models presented in this paper (see columns (1) and (2)). All of the F-statistics are extremely large in magnitude, suggesting that age or season of birth (or both) has a highly significant independent effect on test performance. Season of birth appears to matter most for whites, for which there is a strong first-stage relationship between quarter of birth and schooling in the summer of 1980 (F-statistic=37.3). This wrongly suggests that school entry legislation is a strong predictor of schooling, even though we uncover no strong evidence that there is discontinuity of completed schooling at school entry cutoff dates for whites in the NLSY. A similar pattern of first-stage estimates results when the full set of controls is included in the

---

[39] Neal and Johnson (1996) use a sample of respondents from the 1962 to 1964 birth cohorts of the NLSY. They report an effect of schooling on the AFQT of 0.22 standard deviations for men and of 0.25 standard deviations for women. For the specification in column (3), we estimate an effect of 0.219 for the pooled sample (including race fixed effects), when limited to respondents born after 1961 (n=3126). When we use our entire sample (n=3300), we estimate an effect of 0.2 for the pooled sample.

model (column (4)), even though the TSLS results are more comparable to those presented in this paper.

This comparison suggests that it is indeed inappropriate to instrument for schooling using quarter of birth in this application. As alluded to above, quarter of birth by itself likely captures some combination of the effects of school entry legislation, a weak seasonality in fixed family-related characteristics, and the cumulative effects of investments made outside of school, each of which should be correlated with schooling.

Nonetheless, two caveats about our results are worth noting. First, our estimates capture the combined effects of schooling and age at school entry. Returning to the stylized "local" comparison with a single cutoff in (3) above, we are in essence comparing those who are the oldest in one grade to those who are youngest in the next. In light of evidence that being young in one's grade negatively effects test performance (Datar, forthcoming), our estimates may be biased downward. However, age controls partially capture the effect of age within grade.[40] This might also explain why estimates of the schooling effect tend to rise as the age controls become less restrictive (Table 6).[41]

A second caveat is that our estimates identify the effects of schooling for individuals born near cutoff dates and whose school entry decisions are affected by minimum age requirements. "Compliers" may be a positively selected subgroup, since the most common form of noncompliance with school entry laws is to be below grade, a fact that should be kept in mind in

---

[40] For example, children born in July and August will always tend to be among the youngest in their class, while those born in January and February will always tend to be among the eldest. Whether children in September through December are relatively young or relatively old depends upon the relevant school entry cutoff date.

[41] The first stage effects of academic cohort on schooling also diminish in these specifications. This would be expected if being young in one's grade also increases the probability of grade retention and delayed school entry. For example, in the 1993 and 1995 National Household Education Surveys, between 11 and 13 percent of first and second graders born in the third and fourth quarters of the year experienced delayed entry into kindergarten, compared to only 5 to 6 percent of children born earlier in the year. Children born in the second half of the year were also more likely to have been retained in kindergarten (National Center for Education Statistics, 2000). Grade retention has also been directly shown to be correlated with relative age in other work (Eide and Showalter, 2001).

comparing our estimates to those from other studies.[42] Unfortunately, we can say little beyond

this, since our sample is too small to learn much about how the effect of schooling may vary by

observable characteristics, though our estimates are consistent with a larger impact for minorities

than for whites.

VII.   Conclusion

Ever since the Armed Forces Qualifying Test was first administered by the U.S. military

during World War I, the relative importance of "nurture" in explaining test performance has been

an issue perennially revisited by academics (Jencks and Phillips, 1998). In recent years, the

AFQT, available in the NLSY 79, has been widely used by researchers as a measure of skills that

would otherwise be unobservable. Despite years of research, the question of whether these skills

are largely innate or learned—and therefore amenable to investments in human capital—remains

an open one. Credible evidence of the role of nurture that should be critical to interpreting

results in such studies, as well as to determining whether the AFQT is in fact appropriate for any

given application.

In this paper, we have subjected the ability interpretation of the AFQT to a rigorous test.

Building on the work of Neal and Johnson (1996), we have presented new evidence on the

impact of schooling on the AFQT for the youngest respondents of in the NLSY 79. Using

information on exact birth dates, and matching school entry to individuals on the basis of likely

state of residence at age six, we have identified the academic cohorts of each NLSY respondent.

Exploiting variation in cutoff dates across states, we then used academic cohort as an instrument

---

[42] The TSLS estimates therefore have a "local average treatment effect" interpretation (Imbens and Angrist, 1994; Angrist, Imbens, and Rubin, 1996).

for schooling in a model that allows for relatively unrestricted direct effects of birth date—and implicitly, age—on test performance.

Our results provide evidence in support of an achievement interpretation for the AFQT. Although our research design, combined with the NLSY sample, does not have the power to detect an effect for whites, we do find that an additional year of high school has large effects on the AFQT performance of minorities. The TSLS estimates for blacks and Hispanics are not distinguishable from their OLS counterparts and are larger than other estimates in the literature for pooled samples (e.g., Neal and Johnson, 1996; Winship and Korenman, 1997; Hansen, Heckman, and Mullen, 2004). Our results are perhaps most convincing for blacks, where we find no evidence that academic cohort is directly related to other determinants of test scores. We estimate that a one-year increase in schooling at the test date raised AFQT performance by around 0.35 standard deviations—nearly one-third of the black-white gap in average test performance.

Although in practice the black-white test score gap remains roughly constant for the cohorts under consideration, these results suggest that schools may play an important role in keeping the black-white test score gap from dramatically expanding as children age. Further research will be required to disentangle the relative importance of schooling and family background to the evolution of test performance over the lifetime.

References

Angrist, Joshua, Guido Imbens, and Donald Rubin (1996). "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434): 444-55.

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

---- (1992). "The Effect of Age at School Entry on Educational Attainment: An Application on Instrumental Variables With Moments From Two Samples," *Journal of the American Statistical Association* 87(418): 328-336.

Bock, R. Darrell and Elsie G. Moore (1986). *Advantage and Disadvantage: A Profile of American Youth.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bound, John and David A. Jaeger (2000). "Do Compulsory School Attendance Laws Alone Explain the Association Between Quarter of Birth and Earnings?" in Polachek, Solomon ed. "Worker Well-Being" in *Research in Labor Economics* Vol. 19 Amsterdam: Elsevier.

Bound, John, David Jaeger, and Regina Baker (1995). "Problems with Instrumental Variables Estimation When the Correlation Between and Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* 90(430): 443-450.

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

Cascio, Elizabeth (2004). "Schooling Attainment and the Introduction of Kindergarten Programs into Public Schools." Manuscript, University of California Davis.

Ceci, Stephen J. (1991). "How Much Does Schooling Influence General Intelligence and its Cognitive Components? A Reassessment of the Evidence." *Developmental Psychology* 27: 703-722.

Center for Human Resource Research (2001). *NLSY79 User's Guide: A Guide to the 1979-2000 National Longitudinal Survey of Youth Data.* Columbus, OH: Ohio State University.

Cook, Thomas D. and Donald T. Campbell (1979). *Quasi-Experimentation: Design & Analysis Issues for Field Settings.* Chicago: Rand McNally.

Datar, Ashlesha (forthcoming). "Does Delaying Kindergarten Entrance Give Children a Head Start?" *Economics of Education Review.*

Eide, Eric R. and Mark H. Showalter (2001). "The Effect of Grade Retention on Educational and Labor Market Outcomes." *Economics of Education Review* 20: 563-576

Fischer, Claude S., Michael Hout, Martin Sanchez Jankowski, Samuel R. Lucas, Ann Swidler, and Kim Voss (1996). *Inequality By Design: Cracking the Bell Curve Myth.* Princeton: Princeton University Press.

Fryer, Roland G. and Steven D. Levitt (2004). "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86(2): 447-464.

Hansen, Karsten, James J. Heckman and Kathleen J. Mullen (2004). "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121(1-2): 39-98.

Heckman, James J. (1995). "Lessons from the Bell Curve." *Journal of Political Economy* 103(5): 1091-1120.

Herrnstein, Richard and Charles Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Simon and Schuster.

Imbens, Guido and Joshua Angrist (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467-75.

Jencks, Christopher and Meredith Phillips (1998). "The Black-White Test Score Gap: An Introduction," in Christopher Jencks and Meredith Phillips (eds.) *The Black-White Test Score Gap*, (pp. 1-51). Washington, D.C.: The Brookings Insitution.

Kestenbaum, Bert (1987). "Seasonality of Birth: Two Findings from the Decennial Census." *Social Biology* 34(3-4): 244-248.

Lam, David A. and Miron, Jeffrey A. (1991). "Seasonality of Births in Human Populations." *Social Biology* 38(1-2): 51-78.

McCrary, Justin and Heather Royer (2003). "Does Maternal Education Impact Infant Health? A Regression Discontinuity Approach Based on School Age Entry Laws." Manuscript, University of Michigan and University of California Berkeley.

Murnane, Richard J., John B. Willett, and Frank Levy (1995). "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77(2): 251-266.

National Center for Education Statistics (2000). "Children Who Enter Kindergarten Late or Repeat Kindergarten: Their Characteristics and Later School Performance."

Neal, Derek and William R. Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104(5): 869-895.

NLS User Services (1992). *NLSY79 Profiles of American Youth: Addendum to Attachment 106.*

Seiver, Daniel A. (1985). "Trend and Variation in the Seasonality of U.S. Fertility, 1947-1976."
    Demography 22(1): 89-100.

Todd, Petra E. and Kenneth I. Wolpin (2004). "The Production of Cognitive Achievement in
    Children: Home, School and Racial Test Score Gaps." Manuscript, University of
    Pennsylvania.

Warren, Charles W. and Carl W. Tyler (1979). "Social Status and Season of Birth: A Study of a
    Metropolitan Area in the Southeastern United States." *Social Biology* 26(4): 275-288.

Winship, Christopher and Sanders Korenman (1997). "Does Staying in School Make You
    Smarter? The Effect of Education on IQ in *The Bell Curve*," in Devlin, B., Fienberg, S.E.,
    Resnick, D.P., and Roeder, K. (eds.), *Intelligence, Genes, and Success: Scientists Respond to The
    Bell Curve.* (pp. 215-234). New York: Springer-Verlag.

Appendix A. Coding of School Entry Ages

This paper required knowledge of exact school entry cutoff dates facing several cohorts in the NLSY. We derived dates listed in the Table 1 from archival work based on the histories of current age at school entry statutes and our knowledge of the first grade entry ages in 1955, given in Angrist and Krueger (1992). More specifically, if the current statute's history did not indicate any change in the school entry law over 1955 to 1970, the date recorded pertains to that effective in 1955. On the other hand, if the statute's history indicated a change in the statute over the period, we investigated relevant state session laws over this period. If the state changed its cutoff date, and this cutoff date affected NLSY cohorts, we record the dates in effect in 1968, 1969, and 1970. We later discovered cutoff dates for 1965 and 1972 for a subset of states in the *Digest of Education Statistics*; dates that we recorded are largely consistent with those recorded from this source.

Appendix B. The NLSY Sample

	Number of observations dropped	Total
Number of NLSY 79 respondents (Base year Survey)		12686
Drop if:		
Did not take AFQT/ASVAB	772	11914
Cannot impute schooling for summer 1980	24	11890
Not born in U.S.	778	11112
Born in U.S. or in Puerto Rico or outlying area	130	10982
State of birth missing, state of residence at age 14 missing	3	10979
State of birth missing, have moved since birth	35	10944
AFQT/ASVAB completed under altered test conditions	21	10923
Missing day of birth	15	10908
States where cutoff is at discretion of LEA (GA, IN, MA, WA)	985	9923
1962 - 67 academic cohorts	6178	3745
1971 academic cohort	54	3691
Supplemental sample, white	391	3300
Size of sample used in this study		3300

TABLE 1 - CLASSIFICATION OF STATES BY MINIMUM AGE AT SCHOOL ENTRY LAWS, 1968-1970

	States
Third Quarter Cutoffs	
September 1	CO, DE$_{(68)}$, KS, MD, MN, MT, TX
September 2	UT
September 13	NH
September 15	IA, WY
September 30	MO, TN, VA
Fourth Quarter Cutoffs	
October 1	AL, AR, NJ, NC$_{(68,69)}$
October 15	ME, NE, NC$_{(70)}$
October 16	ID
October 31	ND, OH, SD
November 1	OK, SC, WV
November 2	AK
November 15	OR
December 1	CA, IL, MI, NY
December 31	KY, LA, HI, NV, RI, WI
First Quarter Cutoffs	
January 1	AZ, CT, DE$_{(69,70)}$, FL, MS, NM, VT
February 1	PA
At Discretion of Local Education Agency	
	GA, IN, MA, WA

Notes: A child is permitted to begin first grade in a given academic year if that child reaches the age of six by the date specified. Dates are relevant for the youngest of the NLSY cohorts, eligible to enter first grade between the fall of 1968 and the fall of 1970. See Appendix A and text for more details.

TABLE 2 - ILLUSTRATION OF THE IDENTIFICATION STRATEGY:
BLACKS IN ESTIMATION SAMPLE

	Regression Discontinuity		Differences-in-Differences	
	Schooling, 1980 (1)	Revised AFQT (Standardized) (2)	Schooling, 1980 (3)	Revised AFQT (Standardized) (4)
A. Born in August or September				
Born in September			-0.135 (0.177)	0.025 (0.156)
Sept 1 Cutoff			0.377 (0.458)	-0.260 (0.404)
Born in Sept *Sept 1 Cutoff	-1.445 (0.773)	-0.187 (0.243)	-1.198 (0.568)	-0.053 (0.500)
TSLS Effect of Schooling, 1980 on AFQT		*0.129 (0.172)*		*0.045 (0.410)*
B. Born in September or October				
Born in October			0.092 (0.220)	-0.140 (0.162)
Sept 31 or Oct 1 Cutoff			0.223 (0.311)	-0.104 (0.228)
Born in Oct *Sept 31 or Oct 1 Cutoff	-0.721 (0.393)	-0.230 (0.279)	-0.830 (0.479)	-0.132 (0.352)
TSLS Effect of Schooling, 1980 on AFQT		*0.319 (0.364)*		*0.159 (0.390)*
C. Born in November or December				
Born in December			-0.027 (0.201)	-0.089 (0.170)
Dec 1 Cutoff			0.925 (0.329)	0.483 (0.278)
Born in Dec *Dec 1 Cutoff	-0.855 (0.222)	-0.023 (0.391)	-0.907 (0.425)	0.126 (0.358)
TSLS Effect of Schooling, 1980 on AFQT		*0.027 (0.455)*		*-0.139 (0.425)*
D. Born in December or January				
Born in January			-0.142 (0.161)	-0.001 (0.139)
Dec 31 or Jan 1 Cutoff			0.771 (0.282)	0.069 (0.243)
Born in Jan *Dec 31 or Jan 1 Cutoff	-0.619 (0.337)	-0.143 (0.324)	-0.480 (0.384)	-0.149 (0.330)
TSLS Effect of Schooling, 1980 on AFQT		*0.230 (0.478)*		*0.309 (0.610)*

Notes: The underlying sample includes all blacks in the cross-sectional and supplemental NLSY samples. The dependent variable is either schooling completed as of the summer of 1980 (columns (1) and (3)) or the revised AFQT score, standardized to have a mean of zero and a standard deviation of one (columns (2) and (4)). Standard errors are given in parentheses. Regressions in Panels A, B, and C are based on blacks born in one of the months specified during either 1962 and 1963. Rregressions in Panel D are based on blacks born in December 1962 or 1963 or January 1963 or 1964. All regressions include "comparison fixed effects" (year of birth fixed effects in Panels A to C and an indicator for born in December 1963 or January 1964 in Panel D), and are weighted by AFQT sampling weights. Regressions given in columns (1) and (2) are based on 14, 32, 27, and 29 observations, respectively. Regressions given in columns (3) and (4) are based on 143, 141, 131, and 161 observations, respectively.

TABLE 3 - DESCRIPTIVE STATISTICS FOR THE NLSY SAMPLE

	All		White		Black		Hispanic	
	Mean (sd)	F-Stat	Mean (sd)	F-Stat	Mean (sd)	F-Stat	Mean (sd)	F-Stat
A. Schooling and AFQT Measures:								
Schooling, 1980	10.64 (1.13)	10.7 ***	10.70 (1.10)	5.1 **	10.47 (1.17)	22.2 ***	10.31 (1.26)	2.7
Age as of July 1, 1980 (days)	17.14 (0.87)	0.0	17.14 (0.86)	0.0	17.14 (0.88)	0.1	17.11 (0.87)	1.1
AFQT Score (Percentile)	44.08 (27.82)	0.0	49.42 (26.85)	0.0	20.50 (18.81)	4.0 *	30.19 (23.96)	0.5
AFQT Score (Standardized)	-0.24 (1.00)	0.2	-0.04 (0.92)	0.2	-1.14 (0.81)	3.3 *	-0.74 (0.92)	0.2
ASVAB Arith. Reasoning	-0.20 (0.95)	0.0	-0.02 (0.93)	0.0	-0.99 (0.63)	2.1	-0.63 (0.80)	0.4
ASVAB Word Knowledge	-0.31 (0.98)	0.5	-0.11 (0.87)	1.0	-1.22 (0.97)	0.3	-0.79 (1.01)	1.6
ASVAB Paragraph Comp.	-0.24 (1.07)	0.7	-0.05 (0.99)	0.8	-1.08 (1.04)	5.7 **	-0.70 (1.06)	0.4
ASVAB Math Knowledge	-0.07 (0.99)	0.0	0.08 (0.99)	0.3	-0.72 (0.71)	11.7 ***	-0.49 (0.87)	0.1
B. Background Variables:								
Dad's Highest Grade	11.93 (3.22)	0.1	12.34 (3.08)	0.9	10.57 (2.91)	1.1	9.61 (3.91)	1.0
Mom's Highest Grade	11.72 (2.51)	2.8	12.04 (2.26)	5.0 **	11.04 (2.44)	0.1	8.95 (3.76)	0.3
No. of Children in Family	3.20 (2.19)	0.1	2.91 (1.85)	0.8	4.42 (2.92)	0.0	4.15 (2.91)	2.4
Migrated from State of Birth	0.21 (0.41)	0.0	0.22 (0.41)	0.3	0.18 (0.38)	0.5	0.20 (0.40)	14.0 ***
Enrolled in School, 1979-80	0.91 (0.28)	0.8	0.92 (0.28)	0.7	0.91 (0.28)	0.2	0.87 (0.34)	0.1
Female								
C. School Entry Variables:								
1st Quarter Cutoff	0.12		0.11		0.13		0.11	
3rd Quarter Cutoff	0.25		0.24		0.21		0.37	
4th Quarter Cutoff	0.52		0.66		0.64		0.64	
N	3300		1725		1028		547	

Notes: See text and Appendix B for description of sample selection. The F-statistic tests the significance of the instrument from a regression of the variable specified on academic cohort, state of birth fixed effects, and year of birth by month of birth fixed effects. * Significant at the 0.1 level. ** Significant at the 0.05 level. *** Significant at the 0.01 level.

TABLE 4 - CHARACTERISTICS OF THE 1968-1970 ACADEMIC COHORTS AS
OBSERVED IN THE 1970 AND 1980 DECENNIAL CENSUSES

	White, Non-Hispanic		Black		Hispanic	
	Mean (sd)	F-Stat	Mean (sd)	F-Stat	Mean (sd)	F-Stat
A. Characteristics in 1970						
Highest Grade Attended	1.16	10.5 ***	1.31	57.9 ***	1.14	0.83
	(0.88)		(0.95)		(0.88)	
Mom's Highest Grade	11.40	1.1	9.54	0.0	9.62	2.3
	(2.52)		(2.96)		(3.50)	
Mom's Children Ever Born	3.51	3.1 *	5.16	0.2	4.34	1.0
	(1.89)		(2.91)		(2.38)	
Migrated from State of Birth	0.19	0.0	0.13	0.0	0.21	5.0 **
	(0.39)		(0.34)		(0.41)	
N	24153		6171		1095	
B. Characteristics in 1980						
Highest Grade Completed	9.68	15.0 ***	9.45	48.9 ***	9.50	0.7
	(1.18)		(1.33)		(1.32)	
Highest Grade Attended	10.62	16.5 ***	10.38	47.7 ***	10.41	0.5
	(1.18)		(1.31)		(1.33)	
Enrolled	0.88	0.1	0.87	4.0 *	0.84	0.3
	(0.32)		(0.33)		(0.37)	
Migrated from State of Birth	0.25	0.1	0.20	0.6	0.23	0.4
	(0.43)		(0.40)		(0.42)	
N	119729		32257		6138	

Notes: Underlying data are from the 1970 Form 2 (15%) state file and the 1980 state (A) file. The 1970 sample represents one percent of the population, and the 1980 sample represents five percent of the population. See text for description of sample selection. The F-statistic tests the significance of the instrument from a regression of the variable specified on academic cohort, state of birth fixed effects, and year of birth by quarter of birth fixed effects. * Significant at the 0.1 level. ** Significant at the 0.05 level. *** Significant at the 0.01 level.

TABLE 5 - REDUCED-FORM ESTIMATES, BY RACE

	Coefficient on Academic Cohort (Standard Error)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Non-Hispanic White Subsample							
Schooling, 1980	-0.919	-0.473	-0.412	-0.280	-0.197	-0.298	-0.210
	(0.017)	(0.088)	(0.093)	(0.127)	(0.130)	(0.132)	(0.140)
F statistic	2793.6	28.8	19.8	4.9	2.3	5.1	2.3
R^2	0.48	0.5	0.56	0.5	0.57	0.52	0.58
AFQT (standardized)	-0.113	-0.072	0.025	-0.004	0.099	-0.046	0.061
	(0.028)	(0.079)	(0.094)	(0.110)	(0.115)	(0.106)	(0.114)
F statistic	16.5	0.8	0.1	0.0	0.7	0.2	0.3
R^2	0.07	0.07	0.26	0.08	0.27	0.1	0.29
N	1725	1725	1725	1725	1725	1725	1725
B. Black Subsample							
Schooling, 1980	-0.868	-0.666	-0.709	-0.792	-0.811	-0.737	-0.753
	(0.042)	(0.130)	(0.129)	(0.177)	(0.178)	(0.157)	(0.149)
F statistic	437.2	26.1	30.3	20.1	20.6	22.2	25.5
R^2	0.39	0.4	0.46	0.4	0.46	0.42	0.47
AFQT (standardized)	-0.114	-0.041	-0.103	-0.179	-0.237	-0.225	-0.268
	(0.037)	(0.082)	(0.102)	(0.096)	(0.120)	(0.123)	(0.143)
F statistic	9.5	0.3	1.0	3.5	3.9	3.3	3.5
R^2	0.06	0.07	0.21	0.09	0.22	0.11	0.24
N	1028	1028	1028	1028	1028	1028	1028
C. Hispanic Subsample							
Schooling, 1980	-0.762	-0.363	-0.446	-0.358	-0.468	-0.320	-0.421
	(0.065)	(0.141)	(0.130)	(0.152)	(0.170)	(0.194)	(0.205)
F statistic	137.0	6.7	11.8	5.5	7.6	2.7	4.2
R^2	0.34	0.35	0.43	0.36	0.44	0.4	0.47
AFQT (standardized)	-0.065	-0.041	-0.124	-0.005	-0.097	-0.066	-0.147
	(0.075)	(0.233)	(0.184)	(0.157)	(0.126)	(0.140)	(0.096)
F statistic	0.8	0.0	0.5	0.0	0.6	0.2	2.3
R^2	0.12	0.13	0.25	0.15	0.15	0.2	0.31
N	547	547	547	547	547	547	547
Type of Age Control[†]	N	Q	Q	Q, M	Q, M	M x Yr	M x Yr
Family Background Controls	N	N	Y	N	Y	N	Y
State of birth dummies	Y	Y	Y	Y	Y	Y	Y

Notes: The dependent variable is either schooling completed as of the summer of 1980 or the revised AFQT score, standardized to have a mean of zero and a standard deviation of one. Family background controls include mother's highest grade completed, father's highest grade completed, number of siblings as of age 14, and whether the respondent is female. Where background variables are missing, they have been computed with subsample specific means. When background variables are included, we include a vector of dummies for imputed missing values. All regressions are weighted by AFQT sampling weights. Standard errors (reported in parentheses) are robust to arbitrary residual correlation for individuals born in the same state.

[†]Age controls: N=none; Q=quartic in age as of 1 July 1980; M=month of birth dummies, Yr=year of birth dummies, M x Yr = interactions.

TABLE 6 - OLS AND TSLS ESTIMATES OF THE EFFECT OF SCHOOLING ON AFQT PERFORMANCE, BY RACE

	Coefficient on Schooling, 1980 (Standard Error)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Non-Hispanic White Sample							
OLS	0.332	0.519	0.409	0.527	0.417	0.530	0.416
	(0.019)	(0.031)	(0.028)	(0.032)	(0.030)	(0.034)	(0.031)
2SLS (Linear)	0.123	0.153	-0.060	0.015	-0.501	0.154	-0.289
	(0.030)	(0.167)	(0.232)	(0.390)	(0.797)	(0.330)	(0.662)
2SLS (Dummies)	0.123	0.169	-0.070	0.125	-0.423	0.138	-0.147
	(0.030)	(0.169)	(0.234)	(0.311)	(0.622)	(0.329)	(0.565)
$\chi^2_{(1)}$ statistic (Over-id)	0.0	0.0	0.0	0.1	0.0	0.3	0.5
B. Black Sample							
OLS	0.284	0.363	0.306	0.365	0.309	0.369	0.310
	(0.024)	(0.027)	(0.028)	(0.026)	(0.028)	(0.025)	(0.026)
2SLS (Linear)	0.131	0.062	0.145	0.226	0.293	0.305	0.356
	(0.041)	(0.122)	(0.142)	(0.135)	(0.166)	(0.186)	(0.217)
2SLS (Dummies)	0.131	0.063	0.150	0.225	0.298	0.346	0.391
	(0.041)	(0.122)	(0.142)	(0.135)	(0.168)	(0.197)	(0.224)
$\chi^2_{(1)}$ statistic (Over-id)	5.5	0.2	0.4	0.2	0.4	2.5	2.9
C. Hispanic Sample							
OLS	0.296	0.363	0.306	0.365	0.309	0.369	0.310
	(0.022)	(0.027)	(0.028)	(0.026)	(0.028)	(0.025)	(0.026)
2SLS (Linear)	0.085	0.114	0.279	0.015	0.207	0.206	0.349
	(0.092)	(0.639)	(0.396)	(0.434)	(0.244)	(0.370)	(0.229)
2SLS (Dummies)	0.096	0.193	0.329	0.134	0.269	0.213	0.336
	(0.091)	(0.608)	(0.364)	(0.360)	(0.189)	(0.316)	(0.218)
$\chi^2_{(1)}$ statistic (Over-id)	4.2	0.2	0.2	0.1	0.1	0.0	0.0
Type of Age Control†	N	Q	Q	Q, M	Q, M	M x Yr	M x Yr
Family Background Controls	N	N	Y	N	Y	N	Y
State of Birth Dummies	Y	Y	Y	Y	Y	Y	Y

Notes: The dependent variable in all models is the the revised AFQT score, standardized to have a mean of zero and a standard deviation of one. The 2SLS estimates use either academic cohort ("Linear") or academic cohort dummies ("Dummies") as instruments for schooling in 1980. See Table 5 for a list of family background controls. All regressions are weighted by AFQT sampling weights. Standard errors (given in parentheses) are robust to arbitrary residual correlation within the same state. The critical value for the overidentification test (from a χ^2 distribution with 1 degree of freedom) is 3.841 at 5% significance level.

†Age controls: N=none, Q=quartic in age as of 1 July 1980, M=month of birth dummies, , Yr=year of birth dummies, M x Yr

TABLE 7 - ESTIMATES FOR NON-MOVER SUBSAMPLE
AND ASSIGNING CUTOFF DATES WITH STATE OF RESIDENCE, BY RACE

	Non-movers				Cutoff Assigned Using State of Residence (Age 14)			
	Schooling, 1980	AFQT Revised Score (Standardized)			Schooling, 1980	AFQT Revised Score (Standardized)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RF	RF	2SLS	OLS	RF	RF	2SLS	OLS
A. Non-Hispanic White Sample								
Academic Cohort	-0.372	0.082			-0.233	0.075		
	(0.131)	(0.119)			(0.118)	(0.104)		
R^2	0.59	0.27			0.57	0.29		
Schooling, 1980			-0.219	0.438			-0.324	0.412
			(0.355)	(0.032)			(0.508)	(0.027)
R^2			0.13	0.39			0.06	0.39
B. Black Sample								
Academic Cohort	-0.725	-0.163			-0.617	-0.141		
	(0.191)	(0.162)			(0.205)	(0.142)		
R^2	0.45	0.22			0.44	0.24		
Schooling, 1980			0.225	0.319			0.229	0.313
			(0.236)	(0.030)			(0.245)	(0.024)
R^2			0.33	0.34			0.34	0.35
C. Hispanic Sample								
Academic Cohort	-0.639	-0.365			-0.452	-0.257		
	(0.270)	(0.151)			(0.189)	(0.135)		
R^2	0.48	0.36			0.46	0.32		
Schooling, 1980			0.572	0.307			0.570	0.307
			(0.205)	(0.038)			(0.178)	(0.036)
R^2			0.38	0.45			0.35	0.42

Notes: The dependent variable is either schooling completed as of the summer of 1980 (columns (1) and (5)) or the revised AFQT score, standardized to have a mean of zero and a standard deviation of one (all remaining columns). Family background variables are as listed in the notes to Table 5. All regressions include fixed effects for state of residence and M x Yr of birth. All regressions are weighted by AFQT sampling weights, and standard errors (given in parentheses) are robust to residual correlation within state. Regressions given in columns (1) to (4) are estimated using 1345, 843, and 430 observations in Panels A, B, and C, respectively. Regressions given in columns (5) to (8) are based on 1717, 1017, and 527 observations, respectively.

TABLE 8 - THE EFFECT OF SCHOOLING ON THE AFQT: A COMPARISON OF
TSLS ESTIMATES WITH ALTERNATIVE INSTRUMENTS

	Dependent Variable: Revised AFQT Score (Standardized)			
	Instrument:			
	Academic Cohort (Linear)		QOB (dummies)	
	(1)	(2)	(3)	(4)
A. *Non-Hispanic White Sample*				
Schooling, 1980 (TSLS)	0.154	-0.289	0.182	0.100
	(0.330)	(0.662)	(0.091)	(0.087)
F-stat (on age controls, TSLS)	13347.2	778.2	2.7	4.0
F-stat (on instruments, first stage)	5.1	2.3	37.3	37.0
F-stat (on instruments, reduced form)	0.2	0.3	2.3	0.8
N	1725	1725	1725	1725
B. *Black Sample*				
Schooling, 1980 (TSLS)	0.305	0.356	0.378	0.326
	(0.186)	(0.217)	(0.159)	(0.133)
F-stat (on age controls, TSLS)	337	11605668	1.7	1.4
F-stat (on instruments, first stage)	22.2	25.5	12.7	14.0
F-stat (on instruments, reduced form)	3.3	3.5	4.6	3.4
N	1028	1028	1028	1028
C. *Hispanic Sample*				
Schooling, 1980 (TSLS)	0.206	0.349	0.582	0.300
	(0.370)	(0.229)	(0.230)	(0.389)
F-stat (on age controls, TSLS)	17109	1835	7.9	8.0
F-stat (on instruments, first stage)	2.7	4.2	1.7	1.1
F-stat (on instruments, reduced form)	0.2	2.3	2.5	0.5
N	547	547	547	547
Type of Age Control	M x Yr	M x Yr	Yr	Yr
Family Background Controls	N	Y	N	Y
State of Birth Dummies	Y	Y	N	Y

Notes: The dependent variable in all models is the revised AFQT score, converted to standard deviation units. All regressions are weighted by AFQT sampling weights, and standard errors are robust to arbitrary residual correlation within the same state. See Table 5 for description of controls for age and family background.
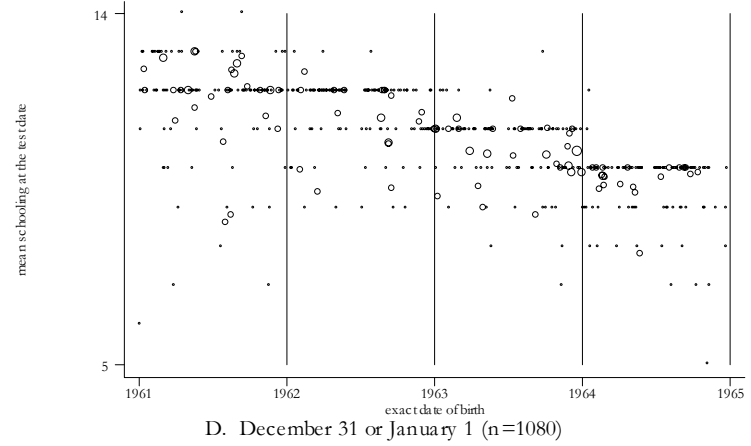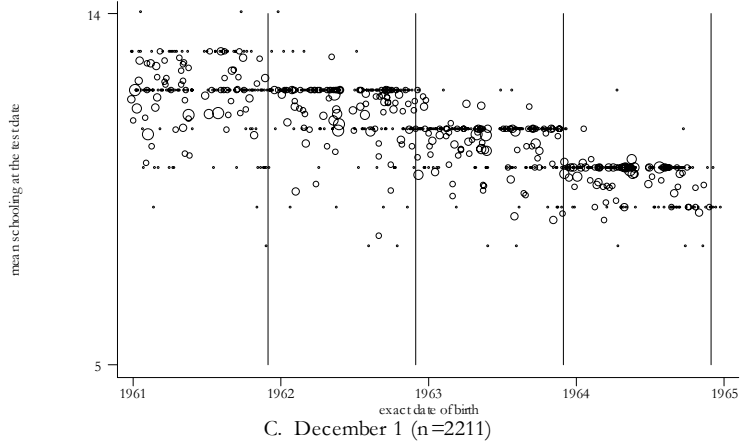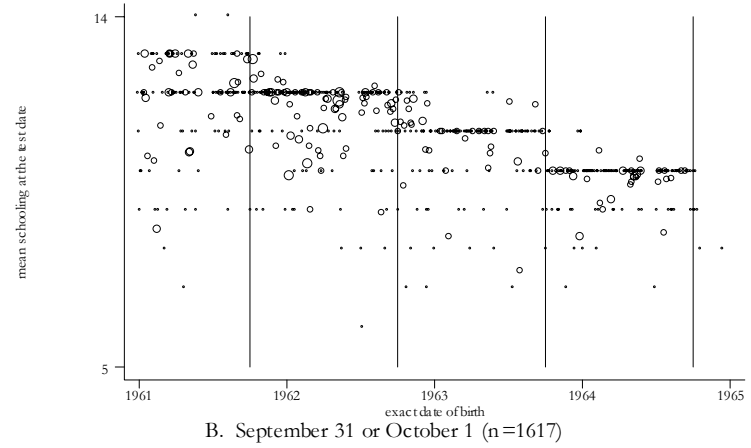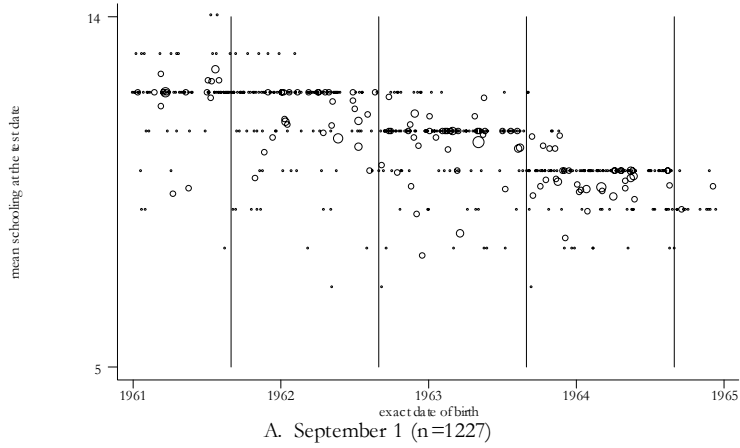
Figure 1: Mean Schooling in 1980, by Cutoff Date

Notes: Sample consists of all native-born NLSY respondents who took the AFQT, and for which schooling in 1980 and exact day of birth are observed. Supplemental samples of blacks and Hispanics are included. Individuals are assigned cutoff dates on the basis of state of birth, using the cutoff date assignment given in Table 1. Averages are weighted by AFQT weights, and point sizes represent the number of observations used to calculate the average.