INTELLIGENCE, GENES, AND SUCCESS

Scientists Respond to The Bell Curve

Bernie Devlin
Stephen E. Fienberg
Daniel P. Resnick
Kathryn Roeder

Editors

COPERNICUS
Blinder defined the following quantities:

\[ E = \sum_{j=1}^{n} \beta_j (X_{ij} - \bar{X}_j) \]  
(A6)

\[ C = \sum_{j=1}^{n} \sum_{i=1}^{n} \beta_j (X_{ij} - \bar{X}_j) \]  
(A7)

\[ U = \beta_0 - \beta_0 \]  
(A8)

\[ D = C + U \]  
(A9)

where \( E \) is the portion of the differential due to differing endowments, \( C \) is the portion of the differential due to differing coefficients, \( U \) is the unexplained portion of the differential, and \( D \) is the portion of the differential attributable to discrimination (see Ref. 9).

References


---

Chapter Ten

Does Staying in School Make You Smarter? The Effect of Education on IQ in The Bell Curve

Christopher Winship and Sanders Korenman

Can education increase an individual's IQ? This has been one of the most incendiary and controversial questions in the social sciences in the past few decades. The greatest firestorm occurred after the publication of Arthur Jensen's 1969 article in *The Harvard Education Review*, "How Much Can We Boost IQ and Scholastic Achievement?" The controversy was further fueled by Richard Herrnstein's 1970 *Atlantic Monthly* article, "IQ." Then, after smoldering for two decades, the passion and acrimony reignited with publication in 1994 of *The Bell Curve* by Richard Herrnstein and Charles Murray.

At the center of the dispute are two related issues: the degree to which intelligence is genetically determined, and thus presumably fixed; and the extent to which observed racial differences in intelligence are genetically based. Jensen, Herrnstein, and Herrnstein and Murray have all been sharply rebuked for suggesting that intelligence and observed racial differences in intelligence might be largely genetically determined.

Education is widely regarded as a key mechanism to elevate the less well-off and to narrow racial differences in social and economic status. A recurring theme in the work of Jensen and Herrnstein has been that although intelligence is a major determinant of economic and social success, because intelligence
is largely genetically determined and fixed, schools will have little success in improving individual economic and social success.

In *The Bell Curve*, Herrnstein and Murray argue and present evidence that education has little to no effect on IQ (Appendix 3, pp. 590 and 591). However, Claude Fischer et al., in a book-length critique of *The Bell Curve*, have argued that the effects of education on IQ are substantial and that the AFQT, the measure of IQ Herrnstein and Murray use in *The Bell Curve*, is primarily a measure of educational achievement.4 The critical question, then, is whether in fact the evidence supports Herrnstein and Murray's assertion of no effect, or is consistent with Fischer et al.'s position.

The purpose of the present chapter is to examine Herrnstein and Murray's analysis of the effects of education on IQ and to review the research literature on this topic. The literature on the effects of educational programs on intelligence is massive. There is, however, surprisingly little consensus as to its findings or their implications. For example, *The Bell Curve* and Stephen Ceci each provide recent extensive reviews but reach sharply different conclusions.4 Herrnstein and Murray are skeptical about the possibility that formal schooling or compensatory programs can increase the intelligence of individuals.5 "Taken together, the story of attempts to raise intelligence is one of high hopes, flamboyant claims, and disappointing results" (p. 389). In contrast, Ceci concludes "that schooling exerts a substantial influence on IQ formation and maintenance...[T]here is now considerable evidence for the importance of variations in schooling on IQ" (see Ref. 5, p. 70).

How are we to explain such divergent summaries of the same literature? A perhaps overly simple answer is that the conclusions reflect political or ideological differences between the scholars. Other factors, however, are important. One critical weakness in the literature has been a lack of focus on determining the size of the key parameter: the effect of schooling on measured intelligence. Herrnstein and Murray as well as Ceci presumably agree that educational programs have some effect on intelligence. The issues in dispute are how much? and under what circumstances? In short, nearly missing from the literature on school effects are sustained attempts to estimate the size of the effect of particular programs. (The literature on Head Start has made such an attempt; see, for example, Ref. 6).

The researchers cited earlier agree that general intelligence, or what psychometricians refer to as g, is a meaningful concept and that it is accurately measured by IQ tests. This position has been contested. Howard Gardner and Robert Sternberg, among others, have argued that intelligence is multidimensional. We do not contribute to this debate (see, however, the chapters in this volume by John Carroll and Earl Hunt). Our interest in IQ begins instead with the considerable predictive power of measured IQ for different dimensions of social and economic success, such as earnings and income,13 job performance,12–14 and criminal behavior.15,16

Herrnstein and Murray's *The Bell Curve* is the most recent and sweeping analysis to show that IQ predicts a variety of behaviors and adult outcomes. Their estimates of the effects of IQ appear to be robust. In an earlier paper we analyzed the sensitivity of Herrnstein and Murray's estimates to different methods for controlling for family background.17 We found that even fixed-effect methods based on sibling pairs, perhaps the most powerful way to control for family background, produces estimates of the effect of IQ on adult behaviors and outcomes quite similar to those reported in *The Bell Curve*. We did, however, find effects of family background comparable in size with those of IQ, and quite a bit larger than those reported in *The Bell Curve*. Also, the effect of IQ was substantially reduced by the inclusion of education as a control. We concluded, as did Jencks (see Ref. 11), that the evidence suggests that IQ—along with education and family background—is an important contributor to social and economic success, but not the dominant determinant, as Herrnstein and Murray stress in *The Bell Curve*. Because IQ is one of the important predictors of success, a logical question to ask is, What determines IQ? In particular, is it true that education has only minimal effects on IQ scores?

The present chapter contains two core sections. In the first we review the small set of articles that have used a methodology similar to Herrnstein and Murray's to estimate an effect of education on IQ. We find that in many studies the estimated effect of education on IQ is larger than that reported by Herrnstein and Murray.

In the second core section, we present a series of analyses of the Herrnstein and Murray model. In the first set of analyses we correct technical problems in the original research involving the treatment of missing data, the omission of a variable for the age at which the initial IQ test was taken, and the fact that some individuals are represented in the data more than one time. These corrections result in an estimated increase of 2.5 points of IQ per year of education, more than double Herrnstein and Murray's estimate.

We next examine four issues in model specification: (1) whether the age at which the AFQT was taken belongs in the model, (2) whether parent's socioeconomic status should be included, (3) whether controls for type of initial IQ test are needed, and (4) whether the appropriate educational variable is educational attainment or years in school. Adjusting the model specification for the first three factors leads to an estimate of 2.3 points of IQ per year of educational attainment. When we use years of schooling as opposed to educational attainment, we obtain an estimate of 1.8 points of IQ for each year of schooling.

Finally, we examine the sensitivity of the results to correcting for measurement error in educational attainment and early IQ. Under what we believe are reasonable assumptions about the extent of measurement error, the estimated effect of education is 2.7 IQ points per year of education. We conclude that,
although it may be impossible to arrive at a single estimate for the effect of education on IQ, a year of education mostly likely increases IQ by somewhere between 2 and 4 points.

The next section discusses our methods and data. The subsequent section reviews the relevant literature. We then present our analysis of the NLSY. We conclude by summarizing what we have learned from the literature review and the data analysis, and we offer some speculative discussion of the implications of The Bell Curve for education policy.

Methods and Data

To conduct our literature search, we began with the recent comprehensive reviews of Ceci and Herrnstein and Murray. We collected the papers cited therein and then, using the Social Science Citation Index, we did forward and backward searches to identify studies that provide estimates (or the data to derive such estimates) of the effects of formal education on IQ. An appendix, which is available from the authors, provides a summary table, essentially an annotated bibliography of these studies.

Our empirical analyses are based on the National Longitudinal Surveys of Youth—the NLSY. For comparability with The Bell Curve, we use data files generously provided by Charles Murray. When necessary we supplement the Herrnstein and Murray data with additional information from the NLSY.

The NLSY is an ongoing longitudinal study of a national baseline sample of 12,868 individuals aged 14 to 20 years as of January 1, 1979. It contains extensive information on individuals’ labor market, schooling, and family formation histories. The importance of the NLSY for the research reported here is that it contains a high-quality measure of mental ability; in 1980 the Armed Services Vocational Aptitude Battery (ASVAB) was administered to nearly the entire NLSY sample.

Our interest in this chapter is in the possible effects of formal education on general intelligence, or g. Within the psychometric tradition g is assumed to be measured by IQ tests. There is considerable consensus in this tradition that many different tests measure g and as a result should be considered IQ tests. The studies we review use a variety of mental ability tests, all of which would be considered IQ tests by psychometricians.

Herrnstein and Murray use the Armed Forces Qualifying Test (AFQT), an equally weighted composite of four items of the ASVAB, as their measure of IQ. They show that the AFQT has the internal structure of an IQ test and that it correlates highly with traditional IQ tests. They further argue that it is one of the best measures of general intelligence currently available. (See, however, the chapter by John Cawley et al. in this volume, in which they argue that the AFQT is a biased measure of general intelligence.)

Previous Research

Any attempt to estimate the effect of schooling on intelligence must confront the problem that individuals with higher initial intelligence may select, or be selected into, higher levels of schooling. Although many studies have discussed and examined the possible effects of schooling on intelligence, few have dealt with this selection problem in even a minimally adequate manner. We restrict our primary attention in this chapter to studies that handle this selection problem by controlling for earlier measures of IQ, the strategy employed by Herrnstein and Murray. Specifically, on pp. 590–591 of The Bell Curve, Herrnstein and Murray present a simple model for estimating the effects of education on mental ability (Fig. 10.1).

The model in Figure 10.1 is closely related to the analysis of covariance model used to estimate the effects of a treatment with pretest/post-test data when there is a nonequivalent control group. Herrnstein and Murray’s model differs from the standard model in that is education has multiple levels, there is considerable time between the pretest (early IQ) and post-test (AFQT), and the time between the two tests and the age at which they are taken vary across individuals. Later we discuss how these differences affect the model specification.

Nearly all the variation in education in the NLSY in 1980, the year that the AFQT was given, is accounted for by differences between individuals in the number of years of high school and college they have completed. As a result, Herrnstein and Murray’s analyses, and our reanalysis of their data, as well as the analyses of most of the other studies we now discuss, involve an assessment of the impact of these years of schooling on IQ.

We have identified six studies that have used an analysis of covariance strategy to estimate the effects of schooling on IQ. Three of the studies are based on U.S. data; the others are from Scandinavian sources. Table 10.1 provides a summary description of this research. The estimated effects range

![Figure 10.1](image)

Figure 10.1. Herrnstein and Murray’s path model. In this model, the impact of earlier IQ score on later IQ score is magnified by a direct effect and an indirect effect mediated through years of education.
from 1 IQ point per year of education in Jencks (Ref. 10) to 4.2 points per year in Husén and Tuijnman (Ref. 23).

With the exceptions of Jencks (see Ref. 10) and Lund and Thrane, none of the studies in Table 10.1 provides a direct estimate of the effect of an additional year of education on IQ. Even when we derive estimates from the results reported, we now discuss each study and explain how we obtained an estimate of the effect of a year of education on IQ.

Lorge’s study consists of a sample of 131 boys who were tested in the eighth grade in 1921–22 in New York City, and were retested 20 years later. The IQ score in the 1920s was a weighted composite of the Thorndike-McCall Reading Scale and the I.E.R. Arithmetic Test. In 1941 two tests were administered—the Otis Self-Administering Test of Mental Ability (Higher Examination, Form B), and the Thorndike Intelligence Examination for High School Graduates (Form V).

Lorge does not provide an estimate of the effect of education on IQ. However, from the information presented in the article it is possible to construct a covariance matrix for initial test score, educational attainment, and the two tests given in 1941. We used this information to estimate a regression model to predict the two different test scores, yielding estimates of 2.8 IQ points for the Otis and 2.37 IQ points for the Thorndike per year of education. We made no attempt to correct for measurement error in the variables. Our analysis assumes that the standard deviation of IQ in this sample is 15, something we could not determine from Lorge’s analysis. This fact, plus the historical age and small size of the Lorge sample, led us to discount somewhat the value of these estimates.

Harnqvist examines a 10% sample of all Swedes born in 1948 who were given a battery of tests at age 13, just before the beginning of school tracking. About 5,000 young men in the sample were retested 5 years later, in 1963, at the time of military enrollment. Harnqvist finds that education had a substantial effect on IQ. He concludes in a later article that his study indicates that, after controlling for family background and initial IQ, one additional year of school increases IQ by 2 to 3 points.

Jencks constructs a series of path models involving a large number of variables and obtains an effect of 1 IQ point per year of education (see Ref. 10). In his analysis he attempts to account for measurement error in the different variables. Jencks gathers the correlations used to construct the path model from a variety of studies conducted at different times and involving somewhat different samples. He recognizes that pooling of correlations across different samples may be problematic and expresses little confidence in his resulting estimate.

Wolffe uses a similar methodology to Jencks (see Ref. 10). In fact, many of his correlations are derived from the same sources. It is perhaps not surprising that the effect of education on IQ is so small that he finds little effect of the home background on IQ in the 10-year-old boys in public and private schools in a Swedish city.
then, he obtains a similar estimate for the effect of education on IQ, 1.07. The correlations used to estimate his path model are from seven different studies. There are three key correlations for our purposes—that between child and adult IQ, between child IQ and education, and from between and adult IQ. The first two correlations are derived from Duncan, 19 Jencks (Ref. 20), and McGee, 21 The final correlation is from Wechsler. 22 The paper does not report the source of the standard deviation of education, although it is most likely to have come from a series of General Social Surveys. As with Jencks's 1972 estimate, we have little confidence in Wolfe's estimate.

One of the most methodologically advanced studies was conducted by Lund and Thrane. A sample of 7,903 children from a southeastern region of Norway, including Oslo, were given a set of military classification tests at the end of grade seven. Four were conventional IQ tests. The same tests were given to 3,450 of the boys approximately 5 years later, at the time of military induction. The final sample with information on both tests and educational attainment contains 2,485 individuals. Lund and Thrane analyze the data in several ways. One of the methods they use is an analysis of covariance model, done with and without corrections for measurement error in the early IQ score. Without the measurement error corrections they obtain an estimate of from 3 to 3.15 IQ points per year of education. * When they correct for measurement error in early IQ, they obtain estimates of from 2.5 to 2.8 IQ points per year of education. The other methods of analysis they use produce somewhat smaller estimates.

Herrnstein and Murray cite Husén and Tuinman as supporting their contention that education has a minimal effect on IQ (p. 591, footnote 17). However, Husén and Tuinman's summary of their results contradicts Herrnstein and Murray's claim. Specifically, they state that not only child IQ has an effect on schooling outcomes but also that schooling per se has a substantial effect on IQ tests scores. Hence, schools not only confer knowledge and instrumental qualifications but also train and develop students intellectual capacity. (Ref. 23, p. 21)

Husén and Tuinman proceed to discuss how their research and that of others demonstrates that IQ does change from childhood to adulthood and how education appears to be a central factor in this change.

The Husén and Tuinman study is a reanalysis of data originally analyzed by Husén. 23 The sample is a cohort of Swedish males (n = 671) who were tested at 10 years of age in 1938 and retested 10 years later at the time of induction into the military. IQ in childhood was measured using four tests: word opposites, sentence completion, perception of identical figures, and restoring disarranged sentences. Adult IQ was measured by four tests: synonyms, concept discrimination, number series, and matrices. Husén and Tuinman use LISREL to analyze the data. They report an estimate of 0.28 for the effect of education on IQ. Rescaling this estimate (multiplying by 1.5), we get an effect of 0.42 IQ points per year of education.

A problem with the Husén and Tuinman study is that education is scaled so that it is only approximately the case that 1 unit is equivalent to 1 year of education. In particular, individuals who have less than a seventh grade education and individuals who have some university education are collapsed, respectively, into single categories at the bottom and at the top of their scale. The effects of doing this are ambiguous. On the one hand, collapsing a continuous or interval scale variable typically reduces the association between it and other variables. This would result in Husén and Tuinman underestimating the effect of education. On the other hand, individuals in the top and bottom categories of the scale differ by 5 points on Husén and Tuinman's scale, but probably on average differ by more than 5 years in their actual educational attainments. The effect of this, then, is to overstate the effect of education.

A similar problem is that in the case of some of the intermediate categories it is possible that a single category of education can represent different amounts of educational attainment. For example, their category "incomplete secondary education" can represent 1 or 2 years beyond the 7 years of compulsory education. As with the two extreme categories, the effects of collapsing are ambiguous. Given that the Husén and Tuinman's estimate is the highest reported in Table 10.1, it is likely it overstates the effect of education on IQ in their sample, perhaps considerably so.

Four other studies worth mentioning have provided estimates of the effect of education on IQ using other methods. DeGroot 33 and Green et al. 34 use "experiments in nature" to estimate the effects of school on IQ. DeGroot exploits the fact that World War II interrupted the schooling of many children in the Netherlands. He compares the IQs and education of successive cohorts of applicants to a training school. He finds an approximate 3-point drop in IQ among individuals most affected by the war, and a deficit of 1.5 years in the amount of time they have spent in school translating into a 3.33-point loss in IQ per year of school. DeGroot, however, makes no attempt to determine whether the war might affect applicants' IQ through other mechanisms in addition to the suspension of schooling.

Green et al. analyze the impact of the decision by the school board of Prince Edwards County, Virginia to shut down all schools from 1959 to 1964 to avoid racial integration. As a result, many black children received no education, whereas white children attended newly opened private schools.

* These estimates are based on the post-test distribution of scores. This is comparable with our use of the age-standardized AFQT scores. Estimates scaled relative to
Ceci, in summarizing the results of this study, concludes that a 1 year loss of education resulted in the 6-point drop in IQ. Unfortunately, the schools in Prince Edwards County were unwilling to make available earlier IQ scores for children in the study. As a result, estimates are based on a comparison of the IQs of children in Prince Edwards County with similar children in adjacent counties. Although Green et al. go to great lengths to achieve comparability between these samples, it is unclear that they succeed, casting doubt on their unusually high estimate of the effect of schooling.

Cahan and Cohen\textsuperscript{11} and Neal and Johnson\textsuperscript{36} capitalize on the fact that rules about when children are allowed to enter school create a discontinuity in the relationship between age and years of education. For instance, if school policy states that a child can start first grade only if he or she is 6 prior to January 1st, then a child born in December will have one additional year of education compared with a child who is born only one month later in January. Cahan and Cohen use this effect of school policy in a regression discontinuity design. They estimate effects of attending fifth or sixth grade (much earlier years of education than we have been considering) on twelve tests. The range of estimates across these tests is between 1.65 and 7.5 points per year of education, with a mean estimate of 4.4 points per year of education.

Neal and Johnson use the same school policy variation to create an instrument for education based on an individual's quarter of birth. They use the NLSY data restricted to individuals born 1961 or later, a sample approximately double in size of that used by Herrnstein and Murray. Using their instrumental variable estimator, Neal and Johnson obtain an effect of 3.3 IQ points for men and 3.75 IQ points for women per year of education.

Reanalysis of the Herrnstein–Murray Model

Herrnstein and Murray use the fact that for a small portion ($n = 1408$) of the NLSY sample, IQ scores at earlier ages are available from school records, along with a score on the AFQT test and a measure of educational attainment in 1980, the time at which the AFQT was taken. They are thus able to estimate a model of the form in Figure 10.1 for a subsample of the NLSY. This approach is very simple. It might be desirable to have a more detailed model that could capture the dynamic evolution with age of IQ and educational attainment. For the purpose of this chapter, however, we will work with variants of the model in Figure 10.1.

Herrnstein and Murray report estimates from two models. In the first, they estimate the effect of early IQ on later IQ without controlling for education. In the second model they add education. On page 590 they state that their analyses control for years elapsed between when the initial test and the AFQT were taken, the age at the first test, and the particular type of initial test. In both cases their dependent variable is IQ (AFQT score) measured in percentiles. For the sake of comparability, we report results using IQ score in percentiles as the dependent variable. In addition, we present analyses in which IQ is measured in the more traditional form of standard IQ units, where IQ is scaled to have a standard deviation of 15 within each age group. Later we focus on this latter set of estimates, reported in column (3) of Table 10.2.

Herrnstein and Murray’s estimates for the effect of education on AFQT, set of the effects of early IQ, are small. They find that an additional year of education increases an individual's percentile ranking on the AFQT by 2.3 percentage points. They report that when the dependent variable is measured in standardized units, the effect of education on IQ is equal to 1 IQ point per year of education. Model 1 of Table 10.2 reports the effects of education and early IQ taken from our replication of their results.

Technical Corrections

We found it difficult to replicate Herrnstein and Murray’s results. There were two problems. First, Herrnstein and Murray did not include age at which the first test was taken as a control variable in their analysis, contrary to the statement on page 590. Murray (personal communication) stated that it was unnecessary to control for age at first test because both the initial and AFQT test scores were age standardized. We return to this issue later and explain in detail why we disagree with Murray. Here we simply investigate the effects of including age at first test in the model.

The second problem we identified is that Herrnstein and Murray had improperly handled missing data in their original analysis. This was independently discovered by Murray. Their original analysis included seven individuals whose education was coded as 5 instead of as missing. In three cases it is possible to impute values for education in 1980 because these individuals’ education was the same in 1979 and 1981.

A third issue results from the inclusion of individuals who had more than one measure of early IQ. Herrnstein and Murray entered these individuals into the data one time for each separate observation of early IQ, treating each record for an individual as an independent observation (Murray, personal communication). Some individuals contribute two or three observations. The problem in doing ordinary least squares (OLS) in this situation is twofold. First, the OLS formula for standard errors of the slope coefficients is wrong because it fails to take account of correlation across errors for different observations on the same individual. Second, OLS gives too much weight to individuals with several early IQ scores (though all individuals only have one observation on AFQT, the dependent variable) decreasing the efficiency of the estimate.
Table 10.2. Effects of education of early IQ and AFQT for different model specifications: coefficients and standard errors

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>AFQT percentiles</th>
<th>AFQT measured in IQ units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Education Early IQ</td>
<td>Education Early IQ</td>
<td></td>
</tr>
<tr>
<td>Original Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>2.280 0.752 0.110</td>
<td>0.381 0.008</td>
<td>Years elapsed between tests</td>
</tr>
<tr>
<td></td>
<td>0.221 0.015 0.000</td>
<td></td>
<td>Dummy for type of test as controls</td>
</tr>
<tr>
<td>Technical Corrections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>3.195 0.736 1.600</td>
<td>0.372 0.008</td>
<td>Missing data correction</td>
</tr>
<tr>
<td></td>
<td>0.273 0.016 0.014</td>
<td></td>
<td>Model 1 controls</td>
</tr>
<tr>
<td>(3)</td>
<td>4.892 0.711 2.506</td>
<td>0.360 0.008</td>
<td>Model 1 controls + age at first test</td>
</tr>
<tr>
<td></td>
<td>0.371 0.016 0.019</td>
<td></td>
<td>Model 1 controls + age at first test</td>
</tr>
<tr>
<td>(4)</td>
<td>4.746 0.716 2.446</td>
<td>0.360 0.010</td>
<td>Corrects for multiple observations on same individual in data set</td>
</tr>
<tr>
<td></td>
<td>0.401 0.018 0.021</td>
<td></td>
<td>All subsequent models make this correction</td>
</tr>
</tbody>
</table>

Model Respecification

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>AFQT percentiles</th>
<th>AFQT measured in IQ units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>4.746 0.716 2.461</td>
<td>0.360 0.010</td>
<td>Model 1 controls + age at AFQT</td>
</tr>
<tr>
<td></td>
<td>0.401 0.018 0.021</td>
<td></td>
<td>Model 1 controls + age at AFQT</td>
</tr>
<tr>
<td>(6)</td>
<td>4.249 0.664 2.192</td>
<td>0.322 0.011</td>
<td>Model 2 controls + SES</td>
</tr>
<tr>
<td></td>
<td>0.396 0.020 0.027</td>
<td></td>
<td>Model 2 controls + SES</td>
</tr>
<tr>
<td>(7)</td>
<td>4.407 0.655 2.284</td>
<td>0.329 0.011</td>
<td>Model 5 controls without dummies for test type</td>
</tr>
<tr>
<td></td>
<td>0.400 0.020 0.028</td>
<td></td>
<td>Model 5 controls without dummies for test type</td>
</tr>
<tr>
<td>(8)</td>
<td>3.666* 0.684 1.806*</td>
<td>0.344 0.011</td>
<td>*Effect of years in school</td>
</tr>
<tr>
<td></td>
<td>0.412 0.020 0.021</td>
<td></td>
<td>Model 6 controls</td>
</tr>
</tbody>
</table>

Measurement Error Corrections (data collapsed so there is only one observation per individual)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>AFQT percentiles</th>
<th>AFQT measured in IQ units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>6.680 0.892 3.554</td>
<td>0.446 0.012</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.642 0.021 0.354</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
<tr>
<td>(10)</td>
<td>5.128 0.767 2.716</td>
<td>0.385 0.011</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.427 0.020 0.247</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
<tr>
<td>(11)</td>
<td>8.527 0.722 4.469</td>
<td>0.361 0.011</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.752 0.020 0.406</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
<tr>
<td>(12)</td>
<td>4.005 0.936 2.126</td>
<td>0.469 0.011</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.401 0.019 0.223</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
<tr>
<td>(13)</td>
<td>2.860 0.954 1.518</td>
<td>0.479 0.011</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.291 0.019 0.162</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
<tr>
<td>(14)</td>
<td>9.783 0.607 5.096</td>
<td>0.303 0.010</td>
<td>Model 6 controls, reliabilities</td>
</tr>
<tr>
<td></td>
<td>0.810 0.019 0.432</td>
<td></td>
<td>Education = 0; early IQ = 0.8</td>
</tr>
</tbody>
</table>

Early IQ is measured in percentile units. SES, socioeconomic status.

Model 2 in Table 10.2 shows the effect of correcting for missing data. The most recent edition of The Bell Curve (see Ref. 3) provides—without comment—estimates using the corrected data. Examining column 3 we see that the effect of education is increased substantially from 1.1 to 1.6 points of IQ.

It may seem surprising that correcting the missing data on seven individuals in a sample of 1,408 should have such a large effect. The influence of a point, however, is proportional to its distance from the centroid (mean values) of the data distribution. In this sample, the mean education is 11.4. The education of these seven individuals is coded as -5, more than 16 years from the mean. As a result, they have a substantial influence on the estimate of the effect of education.

Model 3 of Table 10.1 reports the effect of education when age at first test is controlled. Note that this is what Herrnstein and Murray state they did in the first edition of their book, but, in fact, did not do. The effect of education is increased further—to 2.5 points of IQ per year of education—more than double the estimate in Herrnstein and Murray's original model.

In order to correct for the fact that approximately 10% of the sample is represented more than once in the data set, we weight each record with a weight equal to the reciprocal of the number of times the given individual appears in the data set. The weights for each individual sum to one. To account for correlated errors we use STATA's Huber procedure for estimating standard errors in the presence of clustered data. This procedure is a generalization of Huber's method (see also Ref. 41) of calculating standard errors with heteroskedastic data. Model 4 provides these estimates. The effect of education on the AFQT is only minimally smaller and still more than twice Herrnstein and Murray's original estimate. Its associated standard error is only slightly larger.

Model Respecification

Models 1, 2, and 3 in Table 10.2 include only a minimal set of controls. In addition, some of the variables are age standardized and some are not. We now consider the effects of respecifying the model in several ways.

As noted, Murray has argued that because the IQ variables in his model are age standardized, there is no reason to include in the model controls for age at either the early test or at the time of the ASVAB (in 1980). The issues here are complicated and are analogous to the problem of detrending time-series data in econometric models. If we have a model of the form:

\[ \text{AFQT} = \beta_0 + (\beta_1 \times \text{early mental ability}) + (\beta_2 \times \text{education}) + (\beta_3 \times \text{age}) + \epsilon, \]

where all the variables are in nonstandardized forms, two issues arise. First, early mental ability should be measured at the same age for all individuals.
in the study. In the NLSY, early IQ is measured across a range of ages. Murray is correct that standardizing early mental ability by the age at which the test taken is an appropriate way to adjust for these differences. Second, it may be desirable to measure AFQT in age-standardized units because we are used to dealing with IQ measures in age-standardized form. If so, and if we want to estimate consistently the coefficients in the above equation, then either all the variables in this equation need to be age standardized with respect to the age at which the AFQT was taken or, equivalently, the age at which the AFQT was taken needs to be included in the model. This is similar to time series analysis, in which one must detrend both the dependent and independent variables in order to obtain consistent estimates.42

Model 5 in Table 10.2 shows the effect of including age at which the AFQT was taken as one of the control variables. As in models 3 and 4, the estimated effect of education on AFQT is more than twice as large as in Herrnstein and Murray's original analysis. In fact, the estimates in model 5 are identical to those in model 4. The reason is that, because age at AFQT = age at first test + years elapsed between tests, controlling for any two of these three variables is equivalent.

In Section III of The Bell Curve Herrnstein and Murray present extended analyses of the effects of the AFQT and parents' socioeconomic status (SES) on a variety of measures of social and economic success. Their measure of SES is an equally weighted combination of mother's and father's education, parental income, and occupational status of the head of household (see Ref. 17 for a discussion of their parental SES measure). Herrnstein and Murray, however, do not control for parents' SES in estimating the effects of education on the AFQT score. We would expect that SES would affect both individuals' educational attainment and their later mental ability. Model 6 in Table 10.2 adds parents' SES as a control. As we would anticipate, the effect of education falls, in this case to 2.2 IQ points per year of education.

Models 1 through 6 use a set of dummy variables to control for the type of IQ test initially taken (as did Herrnstein and Murray's analyses). Students took any of ten tests. In this analysis, as in Herrnstein and Murray's, the IQ tests have been standardized to percentile scores.43 Given this standardization, it is not clear why one should add dummies for the type of test taken. In fact, there is an argument against controlling for type of test. There may be a spurious correlation between the type of test taken and measured IQ. Spurious correlation could arise if, for example, students in one region of the country where IQs are lower were more likely than students in other regions to take one type of test. Model 7 in Table 10.2 estimates the effect of education on IQ, omitting the controls for the type of test. The effect of education is now modestly higher, at 2.5 IQ points per year of education.

All the models we have discussed as well as those in the literature have estimated the effect of educational attainment, not years in school. Because very bright students may skip a grade and less able students may repeat grades, educational attainment is a function of both the number of years one has been in school and the rate of progress through school, which presumably is a function of how bright one is. The question is then whether using educational attainment in 1980 as opposed to years in school in 1980 gives a biased estimate of the effect of schooling on IQ scores. The bias could be positive or negative. On the one hand, educational attainment measures years of schooling with error, and thus the effect of schooling should be downward biased. On the other hand, educational attainment partly reflects early IQ, and because our measure of early IQ is imperfect, early IQ is imperfectly controlled in our model; as a result, the effect of schooling should be upward biased.

We were able to construct a measure of years in school from the NLSY. Because full information on school attendance is not available prior to 1980, and the measure must be constructed from retrospective reports on school attendance collected in different years, our measure is likely to contain substantial measurement error. Model 8 reports the results of using years in school as opposed to educational attainment to estimate the effect of school. The estimate is 1.8 points of IQ per year of schooling, much smaller than our estimate in model 7 of 2.5. We have been unable to determine whether this difference is due to greater measurement error in our years-in-school variable (as compared with the educational attainment variable), or because the effect of educational attainment is biased upward because it is partly a measure of mental ability. We intend to investigate this in future research.

Measurement Error Corrections

Measurement error raises difficult problems. There is strong reason to believe that our measures of both educational attainment and early IQ contain measurement error and that this affects our estimates. Unfortunately, we do not know the reliability (the proportion of variance that is "true" variance relative to the total variance, which consists of true variance and measurement error) of either of these two variables in the NLSY. Work by other researchers is of some help. Examination by Jencks (Ref. 11, Table A2.14) of a variety of different sources produces estimates of the reliability of education that range from 0.54 to 0.93. Work by Orley Ashenfelter and Alan Krueger on twins suggests a reliability of 0.9 for educational attainment. IQ measures typically have reliabilities above 0.9.44
In models 9 through 14 we carry out analyses to gauge the sensitivity of the estimates of the effects of education to different assumptions about measurement error in education and early IQ. For individuals with multiple values on early IQ, we use the mean value of their early IQ and the mean length of time between their tests and the AFQT. This procedure reduces the number of observations in our data set (n = 1253), though it allows us to compute correct standard errors. The estimated effect of education ranges from 5.1 IQ points per year of education when we assume that education has a reliability of 0.8 and early IQ is perfectly measured, to 1.5 IQ points per year of education when we assume that education is perfectly measured and early IQ has a reliability of 0.8. Less extreme assumptions produce intermediate estimates. Our preferred model is 10, where we assume that both education and early IQ have reliabilities of 0.9, consistent with values reported in the literature. Here the estimated effect is 2.7 points of IQ per year of education. Note that both the overall amount of measurement error in the two variables and the relative amounts of error in the two variables affect education’s coefficient. If we assume that both education and early IQ have reliabilities of 0.8 instead of 0.9 (model 9), the estimate of the effect of education is 3.5 instead of 2.7 points of IQ per year of education. If, instead, we assume that the reliability of education is 0.8, but the reliability of early IQ is 0.9 (model 11), then the estimated effect is 4.5 IQ points per year of education.

We also investigated the effects of measurement error in parental SES. As noted, Herrnstein and Murray’s measure of SES is an equally weighted combination of mother’s education, father’s education, occupational status of the head of household, and family income. Jencks reports reliabilities for these variables ranging from .72 to .96 (see Ref. 11). Because SES is a composite of these variables, its reliability will be higher. We estimated models 9 through 14 in Table 10.2 assuming that the reliability of SES was .8 and that it was .9.

*The regression routine that we used to estimate models with measurement error (STATA’s eivreg) does not have an option for computing standard errors with clustered data. This is a problem, given that some individuals appear multiple times in our data set. We have two options. One possibility is to estimate our models using weights, as we have in models (4) through (8), and to report incorrect standard errors. The other option is to collapse the data set so that each individual appears only once in the data set. Specifically, for individuals with multiple values on early IQ, we could use the mean value of their earlier IQ along with the mean length of time between their tests and the AFQT for their values on these independent variables. This produces somewhat less efficient estimates because it reduces the number of observations in our data set (n = 1,253), though it allows us to compute correct standard errors. In Table 10.2 we present estimates using the second approach. Estimates for the effect of education using the first approach are somewhat higher than those shown in column 3 of Table 10.2 for models 9 through 14, though never by more than .15.

When we assumed a reliability of .8 for SES, the estimates for the effects of education in column 3 of Table 10.2 changed minimally, always by less than .1, often substantially less. When we assumed a reliability of .9 for SES, probably a more realistic assumption, the estimates in column 3 changed even less, always by less than .05.

Our analyses here, although more extensive than those of Herrnstein and Murray, are still primitive. We have used only one method to correct for selection in the pretest/post-test design: analysis of covariance. Other methods are available and can produce different results (see Ref. 22). An important issue we have addressed only superficially through age standardization is the fact that early IQ is measured at different ages for different individuals. We have yet to attempt to examine whether the effects of schooling differ by race or ethnicity. We have also made no attempt to investigate the importance of differences in the quality of schooling individuals receive or whether other aspects of family background, aside from SES, might affect IQ. In analyses not presented here, we did examine whether the effects of education might be nonlinear and/or whether attending high school and college might have different effects. Although we found large differences, they were statistically insignificant. As a result, we are unable to draw any clear inference about the differential effect of different levels of education on IQ.

Conclusions

Neither our review of the literature nor our analyses of the NLSY data provide a single number for the effect of education on IQ. If we ignore the most extreme estimates in the literature or in our own analysis, our “best guess” estimate would be somewhere between 2 and 4 points of IQ per year of education. From our own analysis our preferred model is model 10, which assumes reliabilities of .9 for both early IQ and education, and has an estimated effect of 2.7 points of IQ per year of education.

Our analysis to this point leaves open the question of which is the more important determinant of one’s AFQT score—education or early (potentially “inherited”) IQ. As noted in the introduction, Fischer et al. have taken the opposite position from Herrnstein and Murray. Specifically, Fischer et al. have argued that the effects of education on IQ are considerably more important than the effect of early IQ on later IQ. They cite our earlier work, in which we first reanalyzed Herrnstein and Murray’s model as supporting this position. We disagree.

All the models in Table 10.2 suggest that both education and early IQ have important effects on later intelligence as measured by the AFQT. It is difficult, if not impossible, however, to compare unstandardized coefficients. We can get around this problem by standardizing both early IQ and educational
attainment to have standard deviations equal to one. If we assume that each year of education increases an individual's IQ by 2.7 points, and that the coefficient on IQ is 0.36 (the median estimate in models 5 through 14) then the rescaled effect of education is 5.1 and the rescaled effect of early IQ is 10.6, more than twice as large.

Herrnstein and Murray argued in *The Bell Curve* that education has little or no effect on IQ. Through examining the literature and reanalyzing the data, we have shown that this conclusion is not supported. But is the effect large or small? In particular, does this range of estimates suggest that education's effect on IQ is modest, and that education may be quite limited as a policy mechanism for reducing inequality?

Dickens et al. argue that if IQ is as important a determinant of social and economic success as *The Bell Curve* suggests, then investments that increase IQ *even modestly* will have substantial payoffs. It then becomes critical to know how responsive IQ is to education. Dickens et al. point out that Herrnstein and Murray's most conservative estimates suggest that a single point of IQ increases annual earnings by $322. Thus, if a year of school increases IQ by 2.7 points, our preferred estimate, the annual payoff from a year of schooling—through its effect on IQ alone—is $626. Assuming the cost of a year of high school is between $5,000 and $8,000, these estimates suggest a substantial monetary payoff.

Furthermore, Herrnstein and Murray demonstrate that IQ is an important determinant of a variety of dimensions of social and economic success such as reduced involvement in crime. Therefore, education may have a considerable effect, through IQ alone, on these other outcomes. (See Ref. for estimates of the social benefits of reducing criminal involvement: one of the outcomes Herrnstein and Murray examine.) Moreover, the payoffs from additional education discussed so far all work through increasing IQ, and hence do not include the "direct" benefits that education might confer. In our earlier analyses of sibling differences in IQ and education, we found that an additional year of educational attainment raises annual earnings by about $1,300 (compared with one's sibling), controlling for differences in AFQT score.

Ironically, then, if the effect of education on IQ is within the broad range we have estimated, *The Bell Curve*'s demonstration of the importance of IQ for social and economic success (in combination with other evidence of substantial "direct" effects of education) provides evidence for the importance of educational investment as a policy instrument, quite contrary to the conclusion that one might reach from reading *The Bell Curve*.

References


Cognitive Ability, Environmental Factors, and Crime: Predicting Frequent Criminal Activity

LUCINDA A. MANOLAKES

Richard Herrnstein and Charles Murray's *The Bell Curve: Intelligence and Class Structure in American Life* has revived the ongoing debate over the appropriateness and usefulness of IQ as an explanatory variable in models predicting behavior. The book posits that a core human cognitive ability exists and is "one of the most thoroughly demonstrated entities in the behavioral sciences and one of the most powerful for understanding socially significant human variation" (p. 14). The work's main thesis is that an individual's intelligence—no less than 40% and no more than 80% of which is inherited genetically from his or her parents—has more effect than socioeconomic background on future life experiences, including criminal actions.

The authors claim that "high cognitive ability is generally associated with socially desirable behaviors, low cognitive ability with socially undesirable ones" (p. 117). Logistic regression analyses are performed on a subsample of white males of the National Longitudinal Survey of Youth (NLSY), and a relationship between IQ and criminality is found. Herrnstein and Murray find that for both of their measures of criminal activity—self-reported by the respondent or via an interview conducted with an incarcerated respondent between 1979 and 1990—socioeconomic status has an insignificant effect after