Is the Test Score Decline Responsible for the Productivity Growth Decline?

By JOHN H. BISHOP*

The test score decline between 1967 and 1980 was large (about 1.25 grade-level equivalents) and historically unprecedented. New estimates of trends in academic achievement, of the effect of academic achievement on productivity and of trends in the quality of the work force are developed. They imply that if test scores had continued to grow after 1967 at the rate that prevailed in the previous quarter century, labor quality would now be 2.9 percent higher and 1987 GNP $86 billion higher.

Multifactor productivity growth in the nonfarm business sector, which was 2.02 percent per year between 1948 and 1965, slowed to 1.04 percent between 1965 and 1973 and then to 0.2.1 percent per year between 1973 and 1987 (BLS, 1988). If pre-1965 trends had continued, the nation would now be 39 percent richer. The research on the causes of the productivity growth slowdown has examined a long list of potential culprits: rising energy prices, government regulations, a shift of output toward services with low rates of technological progress, short-term managerial horizons, reductions in *R&D*, patents, and innovations, the changing demographic composition of the work force, and declines in work effort. Some of these factors have contributed to the decline, but a large portion of the drop in multifactor productivity growth remains unexplained (Edward Denison, 1984; Martin Baily, 1986; Baily, 1981). The absence of a rebound in multifactor productivity growth during the 1980s is particularly difficult to explain. Despite falling oil prices, lowered marginal tax rates, scaled-back regulation, and the entry of the baby boom generation into their prime working years, multifactor productivity in nonfarm business grew a meager .45 percent per year between 1979 and 1987.

There has been some speculation that the decline of SAT scores may be signaling a large drop in the quality of young entrants into the work force and that this may be responsible for a portion of the productivity growth slowdown (John Kendrick, 1980). Baily examined this issue in 1981 and rejected it as a major cause of the productivity slowdown. He calculated how large the decline in the quality of entering cohorts of labor would have to be to explain one-half of the slowdown in productivity growth between 1968 and 1979. During the last few years of this period, the implied relative quality of entering cohorts would have had to have been 40 percent below the pre-1968 levels. "In my view, such a sharp decline is implausible and of a much larger magnitude than anything implied by the SAT scores or related evidence, which suggests that the cohort-quality hypothesis can at most explain a small fraction of the slowdown." (Baily, 1981, p. 13.)

This paper reexamines this issue and reaches the same conclusion regarding the

*Department of Personnel and Human Resource Studies, New York State School of Industrial and Labor Relations, Cornell University, Ithaca, NY 14851. I am grateful to Robert Forsyth, Daniel Koretz, and Peter Mueser for sharing previously unpublished data with me and to Eric Cousineau for his capable research assistance. Thanks are also due to Koretz, Mueser, Richard Murnane, and two anonymous referees for their helpful comments on an earlier version of the paper. Appendixes presenting the yearly data on test score trends are available by writing to the author.

1960s and the 1970s but a different conclusion regarding the 1980s and 1990s. The timing of the test score decline (starting in 1967 and ending about 1980) is remarkably coincident with the productivity growth decline. But a decline in the academic achievement of teenagers cannot cause a coincident decline in productivity growth, for teenagers receive only slightly more than 2 percent of total wages.

The test score decline started to have important effects on productivity growth during the middle of the 1970s. The significance of the test score decline derives from its large size (1.25 grade-level equivalents) and from the fact that it was a decisive break in a 50-year trend of continuous gains in the knowledge and basic skills of those graduating from high school. If the rate of gain in the academic achievement of new high school graduates that prevailed between 1948 and 1973 had been maintained, workers would now be 2.9 percent more productive. The effected workers will remain in the labor force for 50 years, so the vintage model developed in this paper forecasts even larger reductions in productivity in the coming years. Even with an assumption of big gains in academic achievement in the future, the forecast is for a 6.7 percent labor quality shortfall in the year 2010. The cumulative total social costs through 2010 of the test score decline have a present value (at a 6 percent real discount rate) in 1987 of $3.2 trillion.

The paper is organized in five sections. Section I analyzes Panel Study of Income Dynamics (PSID) data and demonstrates that when the errors in measurement are accounted for the general intellectual ability of an adult has large effects on wages. Section II examines time-series data on academic achievement at specified levels of schooling. Section III develops a vintage-based accounting framework for general intellectual achievement and uses it to describe the impact of the test score decline on work force quality. Section IV of the paper presents estimates of the impact of the test score decline on productivity growth. Section V of the paper discusses the implications of these estimates for historical analysis of productivity growth and for educational and labor market policy.

I. The Impact of General Intellectual Achievement on Productivity

General intellectual achievement (GIA) is a summary term for the developed cognitive abilities, competencies, and knowledge which contribute to productivity in most jobs. Included in this construct are developed abilities such as reading, listening, speaking, writing, analyzing, synthesizing, reasoning, doing mathematics, thinking critically, and knowing important facts and principles of the sciences, history, and art. These abilities are skills essential for performing many job tasks, the tools for learning new tasks, and the foundation upon which much job-specific knowledge is built. Often referred to as academic achievement, the word "intellectual" has been substituted for "academic" in order to bring attention to the fact that much of the learning that generates GIA occurs outside school. In principle, the best measures of GIA are broad-spectrum, general-purpose achievement tests such as the Iowa Test of Educational Development (ITED), the ACT, and the Science Research Associates achievement series.

GIA is also proxied by such familiar "aptitude" tests as the Armed Forces Qualification Test (AFQT), the SAT, the Wechsler Adult Intelligence Scale Revised (WAIS-R) Verbal IQ, the Lorge-Thorndike IQ test, and the G aptitude of the General Aptitude Test Battery (GATB). There is considerable evidence for this assumption: (1) school attendance raises scores on these aptitude tests (Irving Lorge, 1945; Torsten Husén, 1951; Department of Labor, 1970); (2) the environment clearly affects scores as seen in the more than one standard-deviation increase in mean IQ scores of young adults during the postwar period in Japan, France, Germany, and Holland (James Flynn, 1987); (3) for time periods for which comparable data are available, trends of aptitude tests scores appear to parallel trends for achievement tests (Daniel Koretz, 1986), and (4) broad spectrum achievement tests correlate almost as highly with verbal and mathematical apti-

tude tests as alternate forms of the same test correlate with each other.[1] Despite differences in purposes, subject matter, and modes of administration, the tests listed above apparently measure closely related constellations of developed abilities that tend to move together over time. While the assumption that the tests listed above are all indicators of the same latent variable may do some violence to reality, it facilitates a consistent accounting of GIA's effect on productivity growth.[2]

The starting point for such an accounting must be an estimate of the impact of general intellectual achievement on the productivity of individual workers. The standard way to approach this question is to infer the effect of GIA on productivity from its effect on wage rates. Models must be estimated in which wage rates are predicted by a *contemporaneous* measure of GIA while controlling for schooling and other worker characteristics such as experience. It is essential that GIA be measured long after the completion of schooling and as close as possible to the date of the wage rate observation. Studies that have had scores on tests taken many years apart have found that the more recent test is by far the more powerful predictor of earnings (Husén, 1969). The difficulty, however, is that reliable GIA tests are time consuming and costly to administer. Consequently, data sets which measure both adult GIA and earnings for national probability samples are rare. When they are available, the measure of GIA is typically a short form IQ test of rather low reliability. The PSID's measure of GIA, for example, has 13 sentence completion questions (taken from the Lorge-Thorndike intelligence test) and a KR-20 reliability of only .652. The result, of course, is that estimated relationships between GIA and wage rates are attenuated by measurement error.

Consequently, the true impact of GIA and years of schooling on wage rates must be estimated as part of a system of equations that includes a measurement model for GIA, years of schooling (SCH), and family background. Such a model was estimated for 1971 PSID data on male household heads 25- to 64-years-old:

$$(1)\quad \begin{aligned} \text{WEARN} &= a_0 + a_1\text{GIA} + a_2\text{SCH} \\ &\quad + a_3\text{AGE} + a_4\text{NOWHITE} \\ &\quad + \mathbf{a}_5\mathbf{TRUEBG} + u_1 \\ \text{TEST} &= \text{GIA} + u_2 \\ \text{YRED} &= \text{SCH} + u_3 \\ \mathbf{MEASBG} &= \mathbf{c^*TRUEBG} + \mathbf{u}_4 \end{aligned}$$

where WEARN is the log of weekly earnings. **MEASBG** is a vector of imperfectly measured characteristics of the individual's true family background: **TRUEBG** = [FAED, FAOCC (Duncan index), SIBS, FAFOReign, BORNFARM, BORNSOuth]. GIA, SCH, and the elements of **TRUEBG** are latent variables with measurement errors (u_2, u_3, and $\mathbf{u}_4$), which are uncorrelated with each other and with equation error (u_1).

[1]For example, reliabilities for the College Board's afternoon Scholastic Achievement Tests are .90 for English Literature and .87 for Math I and for the morning Scholastic Aptitude Tests are .91–.92. The correlation between Math I and the Math SAT is .83 and the correlation between English Literature and the Verbal SAT is .84 (College Board, 1984, 1987). In contrast, the correlation between math and verbal SATs is .66. There are good reasons for high correlations between past achievement in a subject and scores on aptitude tests designed to predict future achievement in the subject. Past achievement aids learning because the tools (for example, reading and mathematics) and concepts taught early in the curriculum are often essential for learning the material that comes later. Furthermore, aptitude tests are validated on later achievement levels, not on rates of change of achievement. Consequently, many of the items that are included are similar to the items that appear on achievement tests. In recognition of the fact that aptitude test scores are significantly influenced by educational background, the College Board describes the SAT as a measure of "developed verbal and mathematical reasoning abilities." (1987, p. 3)

[2]At some future time it may be possible to extend the analysis by disaggregating GIA into math, verbal, and other kinds of achievement, and adding other competencies such as psychomotor skills to the accounting system. It was not undertaken for this paper because there is no data on trends in psychomotor skills and because different test batteries yield different estimates of trends in math and science achievement.

Var(GIA) is normalized to i, Var(GIA)/Var(TEST) = .652 and Var(SCH)/Var(YRED) = .915. For the three dummy variables (FAFOR, BORNFARM, and BORNSO), reliability is assumed to be .903 and c_i is assumed equal to be .95. For the other three background variables, c_i is assumed to be 1 and the reliabilities are assumed to be .702 for FAED, .735 for FAOCC, and .927 for SIBS (Christopher Jencks et al., 1979, table A2.14).[3] The results for models including and excluding **TRUEBG** are:

$$\text{(2)}\quad \text{WEARN} = \underset{(6.85)}{.204}\text{GIA} + \underset{(9.16)}{.0584}\text{SCH} + \underset{(3.21)}{.004}\text{AGE} - \underset{(.86)}{.04}\text{NOWHITE} + a_0$$

$$R^2 = .255$$

$$N = 1774$$

$$\text{(3)}\quad \text{WEARN} = \underset{(6.26)}{.190}\text{GIA} + \underset{(6.24)}{.0576}\text{SCH} + \underset{(2.92)}{.004}\text{AGE} - \underset{(1.25)}{.06}\text{NOWHITE} + \underset{(.45)}{.005}\text{FAED} - \underset{(1.44)}{.0028}\text{FAOCC} - \underset{(.03)}{.0002}\text{SIBS} + \underset{(1.74)}{.076}\text{FAFOR} - \underset{(3.58)}{.152}\text{BORNFARM} - \underset{(.25)}{.009}\text{BORNSO} + a_0$$

$$R^2 = .268$$

$$N = 1774$$

T-statistics are in parentheses below the coefficient. Except for BORNFARM, none of the indicators of family background have a significant direct effect on weekly earnings. The addition of these variables to the model causes a small (7 percent) reduction in the coefficient on GIA.

If GIA is dropped from model (3), the coefficient on SCH is .0813. Thus adding adult GIA, a major correlate and outcome of schooling, to the model lowers the education coefficient by 29 percent. This implies that the higher levels of GIA that are associated with schooling account for 29 percent of the total effect of schooling on weekly earnings. The large direct effect of schooling even when GIA is measured without error suggests that schooling develops or signals other economically productive talents such as discipline, reliability, perseverance, and occupationally specific skills. It also suggests that employers may not know the GIA of job applicants and employees and may use schooling as a signal of GIA (see John H. Bishop, 1987 for evidence on this).

If there is no correction for errors in measurement in model (3), $a_1 = .109$ and $a_2 = .0596$. This implies that correcting for measurement error increases the estimated effect of GIA by 74 percent and reduces the direct effect of years of schooling very slightly. The accounting exercise that follows adopts .190 as its estimate of the response of the logarithm of the wage to a one population standard-deviation change in adult GIA.[4] These

[3]Comparable analysis for women was not feasible because the IQ test was not given to married women. Hourly wages could not be analyzed because they were not available for much of the sample. The model was estimated in LISREL using a correlation matrix kindly provided by Peter Mueser. Age squared was not included in that correlation matrix because the square term was not significant in Mueser's early runs. Experience and experience squared was not substituted for age because it, inexplicably, results in a negative coefficient on experience.

[4]It is possible that the .190 coefficient is biased upward by the absence of controls for genotype. There are three reasons for believing that if such bias exists, it is limited in magnitude. First, while genotype has probably influenced schooling and adult GIA, it appears to have no direct effect on wages in this data set, for adding the three background variables—FAED, SIBS, and FAOCC—with the highest correlation with genotype IQ did not decrease the coefficient on GIA. It was the addition of BORNFARM and FAFOReign which lowered the GIA coefficient. Second, controlling for family background and genotype by estimating within family models comparing brothers actually increases the effect of GIA relative to cross-section regressions of earnings on education and childhood IQ in Michael Olneck's (1977) Kalamazoo data. Finally, when both

results suggest that if the GIA of people with given levels of schooling changes over historical time, these changes need to be explicitly included in any accounting of changes in labor quality. Evidence that GIA at given levels of schooling has been changing is presented below.

II. Trends in the General Intellectual Achievement of Students at Specified Stages of Schooling

The cultural, economic, and educational environment to which children and adults are exposed has improved dramatically in the last century. If we compare, for example, those born between 1897 and 1901 to those born 50 years later, the proportion born on a farm fell from 42.4 percent to 10.6 percent; the proportion who grew up with only one parent present fell from 17 to 13 percent; the average number of siblings fell from 4.8 to 3.3; and their father's years of schooling rose from 6.9 years to 10.7 years. (Robert Hauser and David Featherman, 1976). Time spent in school has increased dramatically. Between 1890 and 1960, the average length of the school term increased 19 percent and average daily attendance rates rose 40 percent. The number of years of schooling completed increased (by more than 50 percent) but so did the achievement levels of students entering and completing high school. Evidence of improvements in academic achievement comes from examining time-series data on the GIA of students at specified stages in their education.

The Interwar Period. A search was conducted for studies reporting the results of administering the same test to different cohorts of students at the same school. Only a few such studies were available for the interwar years. A study of the school children of eastern Tennessee found that over the decade of the 1930s 1st graders gained 11 IQ points and 7th and 8th graders gained 10.8 IQ points (more than two-thirds of a standard deviation, Lester Wheeler, 1942). A study of two high schools in the midwest found no change in the mean IQ of the students at a small rural high school and a 5-point increase between 1923 and 1942 at a large high school serving a small city and the surrounding county (F. H. Finch, 1946). George R. Johnson (1935) found a 3-point gain in IQ between 1925 and 1935 at Grover Cleveland High School in St. Louis. F. P. Roessel's (1937) comparison of the students in three Minnesota high schools in 1920 and 1934 and E. A. Rundquist's (1936) comparison of Minneapolis high school students in 1929 and 1934 both found increases in mean IQ.

Another method of measuring trends is to collect and analyze local studies reporting mean IQs for an entire class or the entire student body of a high school. If we assume that the random process by which high schools were selected for such studies remained unchanged over time, unbiased estimates of long-term trends can be derived from a sample of such studies.

A summary of 29 large sample (over 500 students) studies covering a total of 130,173 students spanning the period 1917 to 1942 is available in Finch (1946). The studies reported measures of central tendency for an IQ test given to all students in a particular grade or to all students in a high school or group of high schools. Almost all the studies used either the Terman or Otis IQ test. When the mean/median IQs reported in these studies are regressed on time and dummies for the test and grade in school, the following results are obtained:

$$\text{(4)}\quad \underset{}{\text{IQ}} = \underset{(2.03)}{.169}\text{*DATE} + \underset{(1.93)}{2.24}\text{*SENIOR} + \underset{(.21)}{.30}\text{*FRESH} - \underset{(.80)}{.86}\text{*OTIS} + \underset{(.71)}{1.23}\text{*OTHER} \qquad R^2 = .29,\ N = .29$$

Despite major increases in high school attendance, the average IQ of high school stu-

adult and childhood measures of IQ compete, it is the adult test not the childhood test which has by far the biggest effect on labor market success (Husén, 1969).

dents was apparently rising 1.69 IQ points per decade. The population standard deviation (POPSD) of these IQ tests for everyone in an age cohort is 15 points. If the sample is restricted to high school graduates only, the standard deviation is smaller (about 11.12 points). Thus, the regression results imply that during the interwar period the GIA of high school students was rising .0113 POPSDs per year or equivalently .0152 HSGSDs per year. This estimate may well understate the gain for the nation as a whole, because the quality of education was improving more rapidly in the South and in rural areas than in the sample of northern predominantly urban high schools included in the above regression.

The Postwar Rise in General Intellectual Achievement. For the post-World War II era, the best data on trends in the general intellectual achievement of students nearing completion of compulsory schooling come from the Iowa Test of Educational Development (ITED). This data set is extremely valuable because it provides equated data extending back to 1942 and annual data from 1960 to the present (Robert Forsyth, 1987).[5] Because about 95 percent of the public and private schools in the state of Iowa regularly participated in the testing program, the analyses of trends in ITED data for Iowa is not plagued by changing selectivity of the population taking the test. This feature of the data makes ITED trends for Iowa a better representation of national trends prior to 1970 than the ACT, the SAT, and the American Council on Education Psychological Exam. These other tests were at first taken by a highly selected group and only more recently by more representative samples of college-bound students. Trends in scores on these tests are biased by the decreasing selectivity of those who took the test.

Figure 1 plots the trends of 12th grade ITED composite scores and an average of 8th grade scores on the Iowa Test of Basic Skills (ITBS) and 9th grade scores on the ITED (A. N. Hieronymus, E. F. Lindquist, and H. D. Hoover, 1979). Through 1966 the trend was up: at first moderately so, and then dramatically after Sputnik. The rate of gain for this period, .023 HSGSDs per year, was substantially higher than the .0152 HSGSDs per year during the interwar years. When scores are plotted in standard-deviation units, the performance of 12th graders and that of 8th and 9th graders track each other very closely. The gains for 12th graders between 1942 and 1966 are all the more remarkable for they coincide with an increase in the high school graduation rate in Iowa from 65 percent in 1941 to 88 percent in 1968.

Other tests that have been administered for long spans of time to stable test-taking populations also exhibited a positive trend during this period. In Indiana, between 1944 and 1976, 6th graders (adjusted for age effects) gained .576 SDs and 10th graders gained .256 SDs on the Iowa Silent Reading test (R. Farr and B. Tone, 1979). Between 1958 and 1966 Minnesota high school juniors gained .39 SDs on the Minnesota Scholastic Aptitude Test (Edward Swanson, 1973). The periodic national standardizations of the ITED also exhibit an increase during the 1960s.

The Test Score Decline and Partial Rebound. Around 1967 the educational achievement of high school students stopped rising and began a decline that lasted about 13 years. On the ITED the composite scores of Iowa 9th graders dropped .283 SDs and the scores of seniors dropped .35 SDs or about 1.25 grade-level equivalents. Comparable declines occurred throughout the country and for upper elementary and junior high school students as well. From peak to trough the decline for seniors was .38 SDs on the SAT

[5]The ITED was revised six times between 1942 and 1985 to incorporate changes in curriculum. Scores on the new versions of the test were made equivalent to the old by administering the old and new test to large samples of students (in a way that insured random assignment of test version to individual students) and then equating them using the equal percentile method. Note that even though measurement error creates problems when a test score is used as a right-hand side variable, estimates of trends in mean scores are unbiased as long as measurement error is uncorrelated with time.

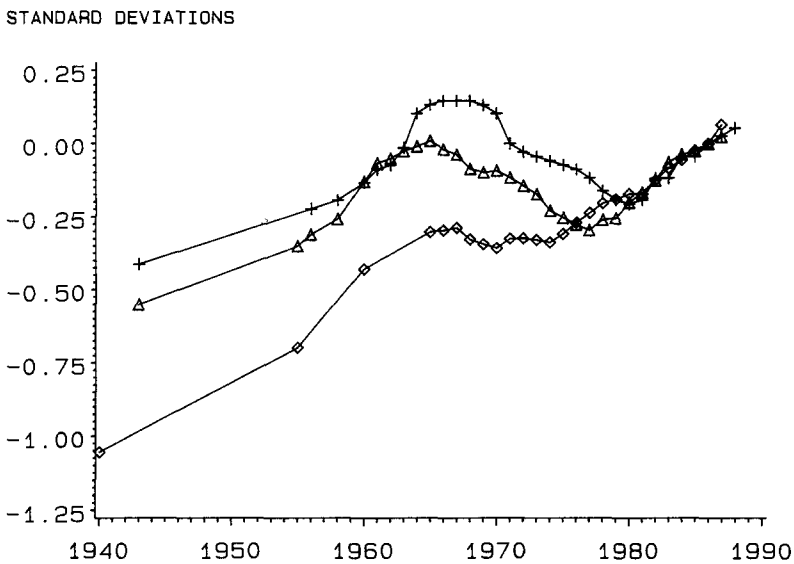

FIGURE 1. IOWA TEST SCORE TRENDS FOR IOWA STUDENTS. + = ITED, GRADE 12. Δ = ITED, GRADE 9 AND ITBS, GRADE 8. ◇ = ITBS, GRADES 3 AND 4. ABSCISSA = STANDARD DEVIATION UNITS; ORDINATE = YEARS. ◇◇◇◇ = GRADES 3 AND 4. ▭▭▭▭ = GRADES 8–9, ++++ = GRADE 12

and .32 SDs on the ACT. For 11th graders it was .28 SDs in the Illinois decade study, .24 SDs on the Preliminary Scholastic Aptitude Test, and .22 SDs on the California Achievement Test. The scores of 9th and 10th graders declined .42 SDs on the Metropolitan Achievement Tests. The scores of 5th through 8th graders declined .33 SDs on the Stanford Achievement Test and .32 SDs on the ITBS (Koretz, 1986; Brian Waters, 1981).

The decline appears to have been caused by something that happened to children after third grade. The IQ of children at entry to school rose .29 SDs between 1964 and 1972. (Flynn, 1984). During the 1970 to 1979 period, when scores of older students were declining, fourth graders in the National Assessment of Educational Progress improved their reading proficiency and comprehension substantially, were stable in mathematics, and declined only slightly in science (Koretz, 1986). Scores of first, second, and third graders on the Iowa Test for Basic Skills both in Iowa and nationally never declined and in fact rose substantially. (See Figure 1). For Iowa third graders there was a gain of .286 SDs per decade (about .7 grade-level equivalents in all) between 1940 and 1966, followed by a period of being stagnant until 1975 when the upward trend of test scores resumed. Since then, Iowa third graders have gained .31 SDs or .3 grade-level equivalents (Hieronymus et al., 1979).

It appears that recent efforts to improve the quality and rigor of the curriculum have had an effect, as test scores are rising again. By 1987 Iowa 12th graders had recouped about 70 percent of their previous decline and 9th and 10th graders had recouped all of a somewhat smaller decline. SAT and ACT scores have risen as well. Rates of gain have been substantial. On the ITED the gain has been .032 SDs per year for 9th and 10th graders. On the ITBS it has been .033 SDs per year for 7th and 8th graders and .027 SDs per year for 3rd and 4th graders. These younger students will not be graduating for four to nine years, so it is very likely that the GIA of high school graduates will continue

to improve for some time. *Nevertheless, general intellectual achievement of high school seniors remains substantially (.554 SDs or 1.96 grade-level equivalents on the ITED) below the level that would have been reached if the trends of the 1940s, 1950s, and early 1960s had continued, rather than reversing after 1966.*

A parallel decline in academic achievement seems to have occurred among those applying to graduate and professional schools. An overall average of scores on the various admission exams taken by this population declined by .121 SDs between 1966 and 1980 (see Appendix A). The decline was no doubt in part due to the substantial increase during this period in the proportion of BA recipients who continued in school beyond the BA. As with high school graduates, there has been a rebound in the test scores (.173 SDs between 1980 and 1986) of college graduates applying to graduate and professional schools.

III. Trends in the GIA of Working Adults with Specified Years of Schooling

The data just reviewed on changes in the GIA of students at specified points in their schooling strongly suggests that, during the first 70 years of the twentieth century, the GIA of the adult population was growing more rapidly than gains in years of schooling alone could account for. Further support for this conclusion comes from comparisons of white enlisted soldiers serving in World Wars I and II. When a stratified random sample of World War II recruits took a test very similar to the test that all literate World War I army recruits had taken, they scored .73 population standard deviations higher (R. M. Yerkes, 1921; R. D. Tuddenham, 1948).[6] The share of this gain due to increases in years of schooling was determined by estimating cross-section regressions of test scores (T) on schooling (S). As one can see in Table 1 an additional year of schooling is associated with a .24 to .25 SDs higher test score in both samples. This means that the two extra years of schooling of the World War II sample explains .48 to .50 SDs of the total .727 SDs test score gain between World Wars I and II leaving a residual improvement of .23 population SDs that must be due to improvements in other environmental factors such as time spent in school per year and the quality of instruction. The tests were taken 25 years apart, so the yearly rate of gain is .0092 POPSDs per year. This result is quite close to the estimate for high school students of .0113 POPSDs per year derived independently from Finch's (1946) data.

A Vintage Model. Estimates of the time path of GIA for working adults were calculated by implementing a simple vintage model of labor force GIA. Since the effects of age and years of schooling are already incorporated in growth-accounting frameworks, what is needed is an estimate of changes in the quality of workers of specified age and schooling which are purged of the effects of changes in the age and schooling composition of the work force. The index to be derived is intended as an additive correction to the labor quality indexes of Dale Jorgensen, Frank Gollop, and Barbara Fraumeni (1987), not as a substitute. Estimates of GIA trends were made for three groups: people who obtain 12 or fewer years of schooling, people who obtain one to four years of college education, and people who enter graduate school. The crucial assumptions of the model are that the *average* GIA of worker cohorts with specified years of schooling are a function of its vintage and its age, and not of the historical epoch during which the aging takes place. Specifically, the mean GIA of any cohort of adults is related to the test scores

[6]The Wells Alpha scores were translated into Army Alpha scores using a table of percentile equivalents developed by Lorge (1936). Means were then calculated and compared to the mean of the World War I army recruits using the SD of the World War I sample for a metric. The resulting estimate of the gain between World War I and World War II is smaller than that reported by Tuddenham (1948).

TABLE 1—CHANGE IN GIA: WORLD WAR I TO WORLD WAR II

	r_{TS}	$\bar{S}$	SDs	B_s	Increase Explained by School	Observed Increase	Residual Due to Other Factors
WWI Sample	.63	8.0	2.6	.242	.484	.727	.243
WWII Sample	.75	10.0	3.0	.25	.500	.727	.227

Source: Tuddenham (1948) provided the raw data on test score distributions for the two populations, the means for schooling, and the correlations between schooling and test scores. The difference between the means of the two samples (.727) was calculated using the table of percentile equivalents between the two versions of the Alpha test provided by Lorge (1936).

this cohort obtained when it was completing schooling by the following equation:[7]

$$(5)\quad \text{GIA(age)} = \text{GIA(17)} + f(\text{age, schooling}).$$

The parameters of $f(*)$ are assumed to have remained unchanged over historical time. For the period after 1942, the key building blocks of the calculation are (1) the time path of scores on tests taken by students entering graduate and professional schools between 1966 and 1986, and (2) the time path of ITED scores for Iowa high school seniors running from 1942 to 1985.

Time paths of mean GIA for those with 0 to 16 years of schooling are derived from the time-series of 12th grade composite ITED scores for 1942 to 1985.[8] The metric of all calculations is a POPSD, the standard deviation of a test given to a random sample of adults. The standard deviation of random samples of high school seniors on these tests is 74 percent of the standard deviation for random samples of adults (U.S. Department of Labor, 1970, Table 20.3). Consequently, all of the ITED trend data discussed in the previous section were multiplied by .74 to translate it into a POPSD metric.

The first step was to construct a time-series of the mean GIA for high school graduates who terminate schooling at the high school diploma. There have been major increases in both the proportion of high school graduates entering college and in the selectivity of college entrance, and adjustments had to be made for these changes.

Paul Taubman and Terence Wales' (1972) careful study of how the selectivity of college

[7]The vintage model does not assume that GIA is constant once a youth leaves school. GIA rises for some and declines for others. The model describes the average GIA of a cohort as maturation and experience produce age-related rises and declines in the mean GIA of the cohort. The time path of mean GIA after completion of schooling is assumed constant over historical time. While this no doubt does some violence to reality, it checks out in the only data sets which provide measures of the GIA of adults that are comparable over historical time —the standardization samples of the Stanford-Binet and Wechsler IQ scales. The IQ of 16–48-year-olds rose .023 POPSDs per year between the 1931/33 and 1953/54 (Flynn, 1984, Table 2) and about .0184 POPSDs per year between 1953 and 1954 and 1976 and 1980. These rates are, as they should be, roughly double the estimated gains in GIA for adults of given years of schooling. It should also be noted that the GIA gain assumed for cohorts of noncollege adults who became 18 prior to 1942 was derived by comparing young adult army recruit samples who had left school an average of ten years prior to testing and then subtracting out the effects of changes in mean years of schooling in the two samples.

[8]Where comparisons are possible, the trends of ITED scores for Iowa seniors are similar to the trends for other tests and elsewhere in the nation. Other states—Minnesota, Indiana, and the national standardization samples of the ITED—also exhibited positive test score trends in the 1950s and through most of the 1960s. From peak to trough, the test score decline was .35 SDs on the ITED, .38 SDs on the SAT, .32 SDs on the ACT. Scores on the ITED are now improving more rapidly than SAT scores, but the slow gains on the SAT appear to be due to significant growth in the proportion of high school graduates who take the test.

TABLE 2—TRENDS IN GENERAL INTELLECTUAL ACHIEVEMENT FOR GIVEN YEARS OF SCHOOL[a]

	1948	1966	1973	1980	1990	2010
GIA 17-Yr.-Olds—*LE* 12 Yrs. of School	.170	.446	.293	.195	.439	.796
GIA 17-Yr.-Olds Entering College	.284	.705	.553	.455	.696	1.055
GIA Shortfall for 17-Yr.-Olds	–	–	(.272)	(.489)	(.419)	(.400)
EQ Index—All Adults	.191	.389	.473	.526	.576	.766
EQ Shortfall—All Adults	–	–	(.014)	(.070)	(.191)	(.352)

[a]Each series is in POPSD units and has been assigned a value of zero in 1929. Figures in parentheses report the difference between the actual/forecasted levels of GIA and what would have happened if the GIA of 17-Yr.-olds had continued to rise at .017 POPSD's per year after 1966. Effects of changes in GIA on the log of labor quality may be obtained by multiplying by .19.

entry has varied over time provides the data for making this adjustment. They report that in 1934 those who continued their schooling after graduation were on average at the 58th percentile in GIA and those who ended their schooling with the diploma were on average at the 43rd percentile. On the assumption that GIA is normally distributed this 15-point difference on a percentile scale at the middle of the distribution corresponds to a GIA differential at entry to college of .385 HSGSDs or .285 POPSDs.

In 1946 college entrants were on average at the 62nd percentile and those not going on to college were at the 43rd percentile. The estimated GIA differential was .39 POPSDs. By 1960 the mean GIA percentile of college entrants in Project Talent data was 62 and the mean percentile of those not going to college was 36. Assuming normality this corresponds to a GIA differential between the two populations of .54 POPSDs. Using these estimates of GIA differentials, an assumption that selectivity of college entrance was constant after 1961, and 1970 census data on the proportion of each cohort which entered college, separate GIA time-series were calculated for those who terminated schooling with the diploma and those who attended college.[9]

The trends of each of the GIA time-series are exhibited in Table 2. The increasing selectivity of college entrance caused the GIA of college entrants to increase more rapidly (.0205 POPSDs per year) between 1942 and 1966 than the GIA of high school graduates who did not enter college (.0136 POPSDs per year). It was assumed that changes in the GIA of those with fewer than 12 years of schooling paralleled the index for high school graduates who ended their schooling with the diploma. The GIA index for college entrants is employed as the GIA index for college dropouts and for college graduates not going to graduate or professional school. Up to 1966, it is also used to index the GIA of those with graduate and professional education. After that date an average of scores on exams for admission to graduate and professional schools is used as the GIA index for this group.

Yearly data on the GIA of school leavers is not available prior to 1942, so a simple trend extrapolation was employed. For those with 0 to 12 years of schooling the GIA trend was assumed to be .0092 POPSDs per

[9]Trends in the mean GIA of college entrants were calculated from the mean (GIA_t) for all high school graduates by $GIA_{ct} = GIA_t + (1-\lambda)(GIA_{ct} - GIA_{ht})$, where λ is the share of an age cohort's high school graduates completing at least one year of college (from 1970 Census data), and ($GIA_{ct} - GIA_{ht}$) is the differential calculated from Taubman and Wales (1972) and GIA_t is a normalized ITED composite score extended back from 1942 at the .0113 SDs per year rate derived from analysis of the Finch data. GIA trends for those who never attended college are obtained from $GIA_{ht} = GIA_t - \lambda(GIA_{ct} - GIA_{ht})$ for the period after 1942.

year. This was taken from the analysis of Tuddenham's (1948) data on army recruits of World Wars I and II presented in Table 1. The GIA trend for those continuing their schooling beyond high school was derived by adjusting the .0113 POPSDs per year trend exhibited in the Finch data for the increasing selectivity of college entrance. Estimates of the selectivity of college entrance are available in Taubman and Wales back as far as 1925. Selectivity was assumed to have been constant prior to 1925. The GIA trend for the 30-year period from 1912 to 1942 was found to be .0136 POPSDs per year.

GIA indexes for the kth age cohort of adult workers with specified years of schooling (GIA_{hkt}, GIA_{ckt}, and GIA_{gkt}) are weighted averages of the GIA indexes for these individuals when they were 17-years-old.[10] The GIA index for 18- to 24-year-olds with less than 13 years of schooling (GIA_{h1t}) is defined as:

$$(6)\quad \mathrm{GIA}_{h1t} = C_{h1} + \left[\sum_{m=t-1}^{t-7} \mathrm{POP}_m \mathrm{GIA}_{hb}\right] \Big/ \sum_{m=t-1}^{t-7} \mathrm{POP}_m.$$

C_{hk} is an appropriately weighted average of f(age, schooling) for the kth age cohort. It is equal to the change in GIA that takes place as a cohort matures from an age of 17 (the age for which test scores are available) to the average age of the kth cohort and is assumed to be constant. The weights (POP_m) are the size of the 1-year age cohort when it was 17-years-old. The GIA indexes for the 25- to 34-year-old cohort and the 35- to 44-year-old cohort were defined as:

$$(7)\quad \mathrm{GIA}_{h2t} = C_{h2} + \left[\sum_{m=t-8}^{t-17} \mathrm{POP}_m \mathrm{GIA}_{hb}\right] \Big/ \sum_{m=t-8}^{t-17} \mathrm{POP}_m,$$

$$(8)\quad \mathrm{GIA}_{h3t} = C_{h3} - C_{h2} + \mathrm{GIA}_{h2t-10}.$$

Figure 2 graphs the cohort specific GIA indexes for those with 12 or fewer years of schooling (assuming constant C_{hk}'s equal to zero). Figures 3 and 4 are the analogous graphs for those with 13 to 16 years of schooling and for those with graduate or professional education.

The vintage model allows us to forecast trends in the GIA of adults well into the 21st century. The only new assumption needed is the rate of GIA gains for entering cohorts. The gains being made in the lower grades right now are quite large, about .020 POPSDs per year, and this rate of gain has been assumed to prevail until 1995. After 1995 it is assumed to drop to .0173 POPSDs per year, the rate of gain that prevailed between 1942 and 1966. This last assumption may be overly optimistic. These projections are displayed in Figures 2, 3, and 4.

The test score decline ripples through the work force, first affecting the productivity of 18- to 24-year-olds, then of 25- to 34-year-olds and so on. At the end of the forecast in 2010, the cohort most affected by the decline will be 45- to 54-years-old and its members will still have 15 years of labor force participation ahead of them. The figures also display the effect of the test score decline on the relative GIA of adjacent age cohorts of workers with the same amount of schooling. Prior to 1973 each generation of new entrants to the work force arrived in their first job better prepared academically than earlier generations with the same amount of schooling. This is no longer the case. This fact is partly responsible for recent declines in the relative wages of youth.

The next step is to calculate for each of the three major categories of completed schooling separate indices of change in the

[10]Since the mean age of those taking graduate school admissions tests is 24, the GIA_{gb} for those who became 24 after 1966 is an average of the graduate admissions scores when the 1-year age cohort was aged 23, 24, and 25. These figures are then averaged to derive GIA_{gkt}. The college graduate and college dropout indexes for 18- to 24-year-olds take into account the fact that the college graduates in this category are almost all age 21 or over and that the college dropouts are mostly over the age of 19.

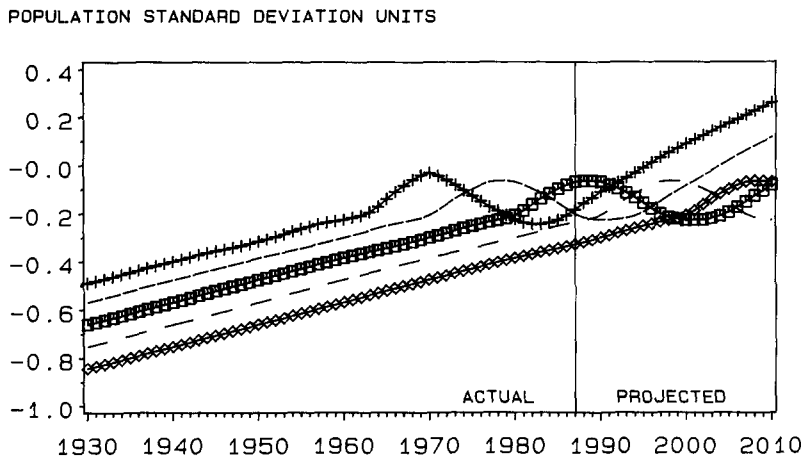

FIGURE 2. GENERAL INTELLECTUAL ACHIEVEMENT TRENDS FOR THOSE WITH 0–12 YEARS OF SCHOOLING BY AGE COHORT. ++++ = 18–24, - - - - - = 25–34, ▭▭▭▭ = 35–44, _ _ = 45–54, ◇◇◇◇ = 55–64

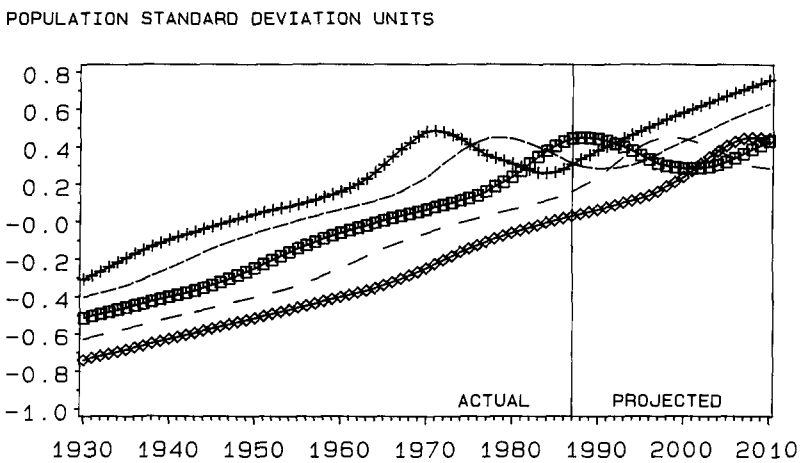

FIGURE 3. GENERAL INTELLECTUAL ACHIEVEMENT TRENDS FOR THOSE WITH 13–16 YEARS OF SCHOOLING BY AGE COHORT. ++++ = 18–24, - - - - - = 25–34, ▭▭▭▭ = 35–44, _ _ = 45–54, ◇◇◇◇ = 55–64

average GIA (holding years of schooling constant) of workers of all ages ($\dot{E}Q_h$, $\dot{E}Q_c$, and $\dot{E}Q_g$). These are chain-weighted averages of the yearly changes in GIA for the six age cohorts indexed by k. For example, the formula for the change in mean GIA of those with fewer than 12 years of schooling ($\dot{E}Q_h$) is:

$$(9) \quad \dot{E}Q_{ht} = \left[\sum_{k=1}^{6} w_{hkt}(\mathrm{GIA}_{hkt} - \mathrm{GIA}_{hkt-1})\right] \Big/ \sum_{k=1}^{6} w_{hkt}.$$

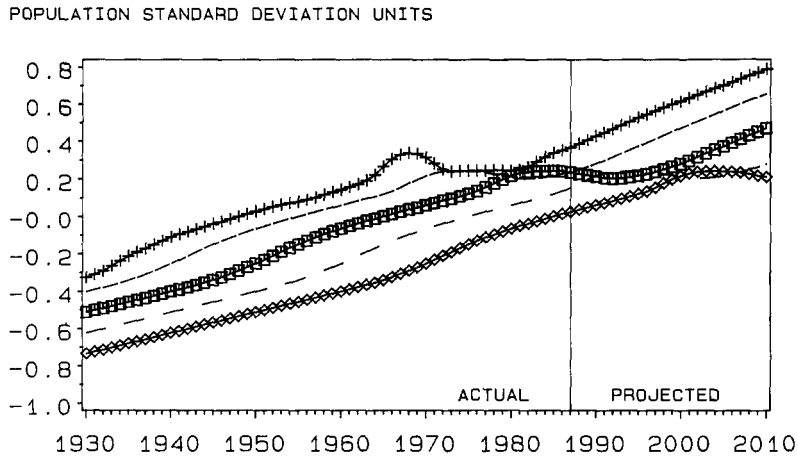

FIGURE 4. GENERAL INTELLECTUAL ACHIEVEMENT TRENDS FOR THOSE WITH GRADUATE EDUCATION BY AGE COHORT. ++++ = 18–24, - - - - - = 25–34, ▤▤▤▤ = 35–44, _ _ = 45–54, ◇◇◇◇ = 55–64

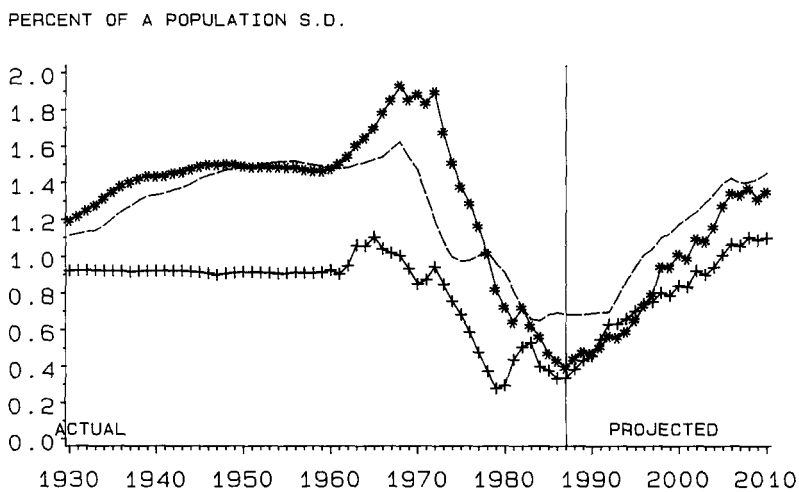

FIGURE 5. RATE OF GAIN IN ACADEMIC ACHIEVEMENT OF THE WORKING POPULATION. SCHOOLING: ++++ = 0–12 YRS., * * * * * * = 13–16 YRS., - - - - - - = MORE THAN 16 YRS.

The weights (w_{hkt}) are an average for the leading and lagging year of the share of total labor compensation going to that age-education group (see Appendix B). While it is tempting to interpret $\dot{E}Q$ as an index of educational quality, this is not really correct for GIA tests measure only one of the objectives of schooling and schooling is only one of many determinants of GIA. Figure 5 provides a plot of the $\dot{E}Q$ indexes multiplied by 100 (i.e., their metric is percentage points of a POPSD of GIA). Because of the increased selectivity of college entrance, the quality of college-educated workers grew faster than

the quality of those with 12 or fewer years of schooling. The aggregate $\dot{E}Q$ index characterizes change in the GIA (holding schooling constant) of the work force averaged over all age-sex-education groups. It is a chain-weighted average of $\dot{E}Q_h$, $\dot{E}Q_c$, and $\dot{E}Q_g$ with compensation shares as weights. $\dot{E}Q$-level indexes were defined by assigning an arbitrary value of zero in 1929 and then cumulating the yearly changes:

$$EQ_t = \sum_{t=29}^{t} \dot{E}Q_t. \tag{10}$$

These indexes are designed to be consistent with Dale Jorgenson (1984), Jorgenson et al. (1987), and Peter Chinloy's (1980) accounting of the effects of changes in the age, sex, and educational composition of the work force on labor quality.[11, 12] The total change in work force quality is the sum of the change in a standard labor quality index ($\dot{L}Q$) and the aggregate $\dot{E}Q$ index (see Table 2) multiplied by the derivative of productivity with respect to test scores scaled in a POPSD metric ($a_1 = .19$):

$$T\dot{L}Q_t = \dot{L}Q_t + a_1\dot{E}Q_t. \tag{11}$$

IV. The Productivity Consequences of the Test Score Decline

Implementation of the GIA accounting system described above produces estimates of the effect on the quality of labor of improvements in the GIA of adults of specified years of schooling. It is estimated that between 1948 and 1973 that gains in the GIA of the working population *not* associated with increases in years of schooling improved labor quality by .212 percent per year or by 5.4 percent over the course of the full 25-year period. The gains in GIA that are associated with increases in years of schooling are roughly equal to those not associated (see fn. 7), so the total effect of GIA gains on labor quality growth is between .40 to .45 percent per year or about 11 percent for the entire 25-year period. This second half of the GIA effect is already included in standard education-based growth-accounting exercises. Jorgenson et al. (1987) estimate that increases in years of

[11]Denison's index of work force quality is not defined in a manner consistent with the EQ index. Unlike Jorgenson and Chinloy, who assume that the full effect of schooling on wages in census tabulations reflects its real value added, Denison adjusts for the higher IQ and SES of those who get additional education. If one desires a full accounting of improvements in labor quality rather than an estimate of the value added of greater amounts of schooling, this downward adjustment is not desirable, however, because parental SES and childhood IQ have been rising and have contributed to improvements in the quality of the labor force. A further source of inconsistency with the EQ index is the inclusion of longer school years and fewer absences in the education quantity series. In the accounting system employed in this paper the EQ index picks up all effects that do not operate through years of schooling. If a rapid expansion of high schools were to lower the average achievement of high school graduates, EQ would decline. Longer school years and lower absenteeism would be expected to raise EQ.

[12]The standard LQ index being referred to is an average of percentage rates of growth of hours worked by groups of workers classified by age, sex, and schooling weighted by shares of compensation. Despite the fact that the weights respond to changes in relative wage rates, changes in quality along dimensions that are not explicitly included in the accounting system do not produce corresponding changes in LQ. For example, if the quality of a particular category of worker falls and this results in an equal percentage decline in the wage of the group, the LQ index will actually be increased if the group at issue grows more slowly than the average for all workers. Jorgenson et al. (1987) also present LQ indexes in which workers have been cross-classified by occupation, industry, and class of worker as well as by age, sex, and schooling. Incorporating these additional quality dimensions adds only 0.12 percent per year to the growth of LQ between 1948 and 1973 and nothing to LQ growth between 1973 and 1979. Some of the LQ change that Jorgenson et al. attribute to changes in these three quality dimensions might in fact be due to changes in EQ. Since academic achievement precedes the choice of occupation, industry, and employer both in time and causally, an accounting framework that incorporates test score changes is in my view preferable. Baily (1986) argues against incorporating occupation, industry, and class of worker in LQ. He points out that transferring a worker from an industry with a low average wage to one with a higher average wage does not (when labor markets are competitive) raise that worker's marginal productivity if the capital employed in the two industries does not change.

schooling caused labor quality to grow .725 percent per year or a total of 19.9 percent over the course of the 25 years.[13] In combination, the gains in years of schooling and in GIA holding schooling constant increased labor quality 26.3 percent by the end of the period.

After 1973, however, gains in years of schooling and the GIA of those completing specified amounts of schooling began to decelerate. Between 1973 and 1979, the contribution of years of schooling to the growth of labor quality diminished to .612 percent per year. The contribution of schooling constant GIA gains to the growth of labor quality reached a postwar peak of .238 percent per year between 1966 and 1973 and then fell to .157 percent per year between 1973 and 1980 and fell even further to .084 percent per year between 1980 and 1987. If, instead, the academic achievement of those leaving school had continued to improve after 1966 at the rate that had prevailed between 1942 and 1966 (.0173 percent of a POPSD per year), gains in GIA holding schooling constant would have increased labor quality by .282 percent per year between 1966 and 1973, by .296 percent per year between 1973 and 1980, and by .324 percent per year in the 1980s.

These findings are consistent with Baily's (1981) analysis quoted at the beginning of the paper. The test score decline was not a contributing cause of the post-1965 productivity growth decline and made only a modest contribution to the post-1973 decline. The test score decline induced reduction in labor quality growth between 1973 and 1980 was .14 percent per year. The contribution to the slowdown in GNP growth and labor productivity during this period is .14 percent per year multiplied by the share of labor in total compensation (about .67) or about .093 percent per year. William Nordhaus (1980) and Denison (1984) have estimated that between 1 and 1.14 percent of the post-1973 slowdown in total factor productivity growth remains unexplained after the effects of energy prices, demographic shifts, the *R&D* slowdown, the growing service sector, and health, safety, and environmental regulations are accounted for. Thus, the test score decline accounts for slightly less than one-tenth of the "unexplained" decline in productivity growth during the 1970s.[14]

The test score decline's major impact on productivity growth has come in the 1980s. During a period in which falling oil prices, lowered marginal tax rates, scaled-back regulations, and an aging work force were expected to cause productivity growth to rebound, the test score decline has been an important drag on productivity growth. The rate of growth of labor quality was .240 percent per year lower between 1980 and 1987 than it would have been if test scores had continued to grow at the rate that prevailed between 1942 and 1966. The drag on productivity growth will continue well into the 21st century. The reduction in labor quality growth resulting from the test score decline is projected to be .19 percent per year in the 1990s and .12 percent per year in the first decade of the 21st century.

The cumulative effect of the test score decline on standards of living is quite large. The labor quality shortfall was 1.3 percent in 1980 and 2.9 percent in 1987. The shortfall is projected to be 3.6 percent in 1990, 5.5 percent in 2000, and 6.7 percent in 2010. The

[13] Estimates of the contribution of schooling to labor quality are obtained by dividing the translog indexes of labor input that take account of age, sex, and schooling by a translog index that accounts for age and sex only. Changes in the age-sex composition of the work force lowered labor quality by 0.16 percent per year between 1948 and 1973 and by 0.58 percent per year between 1973 and 1979. (Jorgenson et al., 1987, Table 8.4).

[14] These estimates of the costs of the slowdown in GIA gains are, of course, sensitive to assumptions made along the way. Their magnitude would be reduced if measurement errors in the Lorge-Thorndike test are assumed smaller or negatively correlated with the true score or if the counterfactual trend in test scores is assumed to be slower. The estimated costs of the decline in the quality of those entering the work force would have been increased if it had been assumed that other productivity enhancing traits such as reliability, good work habits, and vocational skills were declining during the 1970s in much the same way that GIA declined or if as demonstrated in Bishop (1987), difficulties in signaling academic achievement result in its being undercompensated.

social cost of the test score decline is now $86 billion annually (by comparison, total compensation of labor in all public and private schools and colleges was $172 billion in 1986). If the forecasted shortfalls in output up to the year 2010 are cumulated assuming a 3 percent rate of growth of GNP and discounted to 1987 at a real interest rate of 6 percent, the total present discounted costs of the test score decline is $3.2 trillion or roughly three-fourths of the 1987 gross national product.

The direct effects of the test score decline reduced GNP by 0.91 percent in 1980 and 1.9 percent in 1987 and are forecasted to reduce GNP by 3.6 percent in 2000 and by 4.4 percent in 2010 (i.e., roughly two-thirds of the percentage figures quoted for labor quality). When indirect effects of the test score decline are taken into account, the total effect of the GIA decline is likely to be considerably larger than the estimates just quoted. If growth had been higher, the supply of savings and consequently net capital formation would have been higher as well. Additionally, academic achievement appears to enhance the profitability of investments in physical capital and new technology (M. Denny and M. Fuss, 1983; Ann Bartel and Frank Lichtenberg, 1987), so a decline in GIA reduces investment demand and technological progress as well. Only a portion of GIA's effects on the profitability of investment in physical capital and *R&D* turn up in higher wage rates and are therefore included in the accounting (Bishop, 1987), so the estimates of the economic effects of the test score decline just offered are probably conservative.

V. Summary and Implications

The paper has presented evidence that the effect of general intellectual achievement on wage rates and productivity is larger than heretofore believed. It is estimated that holding years in school constant, a one POPSD increase in a worker's true GIA raises productivity by approximately 21 percent (exp .190). This estimate of the effect of GIA is larger than previous estimates for two reasons: wage rates were related to an adult measure of GIA rather than a childhood measure and errors in the measurement of GIA, schooling, and family background were corrected for. This, in turn, implies that the recent test score decline is signaling a significant deterioration in the quality of young entrants into the work force.

The second major finding of the paper is the historically unprecedented nature of the test score decline that began around 1967. Prior to that year, student test scores had been rising steadily for more than 50 years. New estimates of the quality of the work force were developed incorporating the effects of improvements in academic achievement at given levels of schooling as well as increases in years of schooling. Jorgenson et al., (1987) estimate that increases in years of schooling raised labor quality by .725 percent per year between 1948 and 1973. Our estimates imply that improvements in academic achievement at given amounts of schooling contributed an additional .212 percent per year to the growth of the quality of labor during this period. The test score decline reduced this contribution to .16 percent per year between 1973 and 1980, and .085 percent per year in the 1980s. If the test scores of high school graduates had continued to grow at the rate that prevailed between 1942 and 1967, labor quality would now be 2.9 percent higher. The social cost in terms of foregone GNP is now 86 billion dollars annually and it is projected to double within 15 years.

One important implication of the forecasts is that even if current efforts to improve the schools are successful, the test score decline will continue to depress productivity well into the 21st century. Even with rapid improvements in the quality of elementary and secondary education, the labor quality shortfall grows to 5.5 percent in 2000 and 6.7 percent in 2010. The only way to prevent these forecasts from being realized is to change the relationship between GIA at age 17 and GIA as an adult. This might be accomplished by attracting massive numbers of adults back into school, by expanding educational offerings on television and/or by inducing employers to provide general education to long-term employees.

It would appear that the education enterprise has historically been an important source of economic growth. When the academic achievement of students completing their schooling declines substantially, the economic costs are large and last for generations. Consequently, the potential benefits of major improvements in the academic achievement of students would also appear to be substantial. Exactly what caused the test score decline remains somewhat of a mystery. The decline was larger for whites than for minorities and larger in the suburbs than in central city high schools with student bodies from disadvantaged backgrounds (Koretz, 1986). Thus, the educational quality problem is not limited to the schools serving minority and immigrant children. There is no indication that private schools escaped the decline. The test scores of the more able students declined just as much and possibly more than the test scores of less able students. The declines were larger for higher-level skills such as inference and problem solving than for very basic skills such as arithmetic computation (Koretz, 1986). The only place in the educational system where no decline is visible is in the first few years of elementary school. Children arrived in first grade better prepared than earlier generations, they maintained that advantage through third grade, but by the time they graduated from high school in 1980 they had learned about 1.25 grade-level equivalents less than those who graduated in 1967.

Productivity growth and test scores declined almost simultaneously. The changing attitudes toward hard work and authority that are sometimes blamed for the productivity growth decline may have also contributed to the lowering of educational standards. In this sense, the ultimate cause of the test score decline may lie outside the educational system.

International competition forced manufacturing firms to reexamine how they manage and motivate workers. As a result, capital/labor productivity growth in manufacturing has rebounded to 3.6 percent per year between 1981 and 1987, substantially above the 1.6 percent growth rate that prevailed between 1948 and 1981. Competitive pressures cannot, however, be counted on to solve an educational quality problem. It is a systemic problem that is not likely to be remedied without major changes in the recruitment, training, and compensation of teachers, in teaching methods, in the organization of schools, and in how learning is measured and rewarded.

APPENDIX A
Academic Achievement Trends for Those Entering Graduate or Professional Education

The Graduate Record Exam (GRE), the Law School Admissions Test (LSAT), the Graduate Management Admission Test (GMAT), and the Medical School Admissions Test (MCAT) are primarily taken by students applying for admission to graduate and professional schools. Many of those taking these tests are returning to school many years after completing their BA. There have been major changes in the proportions of college graduates continuing their schooling and in the selectivity of the graduate and professional school admissions processes. Consequently, scores on these tests are poor indicators of the GIA of college graduates who do not go to graduate or professional school. They are, however, reasonably good indicators of the GIA at entry into graduate school of those who obtain one or more years of graduate or professional education.

Data on trends for these tests was obtained from Clifford Adelman (1983) and by correspondence with the organizations which administer these exams. The average was designed to characterize the GIA of those with 17+ years of schooling who take jobs in the private business economy. Scores of the 4 different tests were deviated from their value in 1977, divided by their standard deviation, and averaged. The weights for the three professional school tests were the total numbers of test-takers between 1976 and 1986. Because so many of the GRE test-takers do not work in the business sector, the GRE index was given a weight equal to one-half the number of test-takers. The resulting weights were .40 for the GMAT, .22 for the LSAT, .10 for the MCAT, and .29 for an index of GRE subtest scores. The calculation of the GRE index will now be described. Since many of those taking GRE tests are headed for jobs in government and the nonprofit sector, and GRE subject matter scores used in the index were the fields that typically lead to a job in private business: that is, math, biology, physical sciences, engineering, psychology, and economics. Based on the numbers taking the exams between 1976 and 1986 the weights were .27 for biology, .09 for chemistry, .06 for physics, .06 for geology, .06 for math, .14 for engineering, .06 for economics, and .27 for psychology. It was assumed that only one-half of those in economics and psychology enter private business. This GRE Achievement score index was then averaged with the verbal and quantitative GRE "aptitude" test scores. There are breaks in the comparability of the LSAT and MCAT. The 1-year gaps in these series

were filled in by assuming no change in mean scores during the interval.

The academic achievement of those applying for entry into graduate or professional schools also appears to have declined during the late 1960s and 1970s. (See text Figure 1.) Between 1966 and 1977 there was a .13 SDs decline on the quantitative Graduate Record Exam (GRE), a .23 SDs decline on the verbal GRE and a .215 SDs decline in the Graduate Management Admission Test (GMAT). There were small increases of .06 SDs on the Medical School Admissions Test (MCAT) and of .09 SDs on the Law School Admissions Test. An overall average of these scores declined by .107 SDs. The decline is no doubt in part due to the substantial increase during this period in the proportion of BA recipients who entered graduate or professional schools.

As with high school graduates, there appears to have been a rebound in the test scores of college graduates planning to continue their schooling. The overall index fell an additional .014 SD between 1977 and 1980 but has since risen .173 SD. Trends have differed substantially across tests. Between 1977 and 1986 there were declines of .156 SD on the MCAT and .116 SD on the Verbal GRE but increases of .27 SD on the quantitative GRE, .215 SD on the GMAT, and .179 SD on the LSAT.

DATA SOURCE

Adelman, Clifford, *The Standardized Test Scores of College Graduates, 1964–1982*, prepared for the Study Group on the Condition of Excellence in American Higher Education, National Institute of Education, 1983.

APPENDIX B

Construction of the Compensation Shares Used to Weight Age/Education Groups

Compensation share for each age by education group was calculated for 1939, 1949, 1959, and 1969 from the decennial Censuses and Herman Miller (1960). Where 1939 data were not available (as for women), the necessary earnings ratios are derived from the 1950 Census. For 1975 through 1984, the weights are obtained from the table titled Education and Money Earnings in the P60 series of the Current Population Reports. For 1973 and 1974, special total money income tables were obtained from the Census and estimates of unearned income were subtracted. For 1930, the number employed for each age-group was obtained from Series A 119–134 and D 29–41 of Historical Statistics of the United States (1975). The educational background of each age-group was obtained by backward extrapolation from the 1940 Census, and the relative wage rates and employment rates for age/education groups were assumed to be the same as in 1940. For the periods between censuses and for 1971 and 1972, an average of weights at the beginning and end of the time interval are employed. For the 1948 to 1973 period, control totals for compensation paid to those with 12 or fewer years of schooling by sex and 13+ years of schooling by sex were taken from Table 3.8 of Jorgenson (1984). Consequently, only the age breakdowns within educational category remained constant over the intervals between the 1949, 1959, and 1969 censuses. Compensation shares for 1984 were used for the forecasts.

DATA SOURCES

U.S. Bureau of the Census, "Money Income in _____ of Families and Persons in the United States," *Consumer Income*, P60 Series of the Current Population Reports, various issues.

_____, *Census of Population: Detailed Characteristics* for 1980, 1970, 1960, 1950, and 1940, Washington: USGPO.

_____, *Historical Statistics of the United States*, Washington: USGPO, 1977.

Jorgenson, Dale, "The Contribution of Education to U.S. Economic Growth, 1948–73," in Edwin Dean, ed., *Education and Economic Productivity*, Cambridge, MA.: Ballinger, 1984, 96–162.

Miller, Herman, "Annual and Lifetime Income in Relation to Education: 1939–1959," *American Economic Review*, December 1960, *50*, 962–86.

REFERENCES

Baily, Martin Neil, "Productivity and the Services of Capital and Labor," in W. C. Brainard and George L. Perry, eds., *Brookings Papers on Economic Activity*, 1:1981, 1–65.

_____, "What has Happened to Productivity Growth?," *Science*, *234*, October 1986, 443–51.

Bartel, Ann P. and Lichtenberg, Frank R., "Educated Workers and New Technology," *Review of Economics and Statistics*, February 1987, *69*, 1–11.

Bishop, John H., "Information Externalities and the Social Payoff to Academic Achievement," Cornell University: Center for Advanced Human Resource Studies Working Paper No. 87-06, 1987.

Chinloy, Peter, "Sources of Quality Change in Labor Input," *American Economic Review*, March 1980, *70*, 108–19.

Denison, Edward F., *Trends in American Economic Growth, 1929–1982*, Washington: The Brookings Institution, 1984.

Denny, M. and Fuss, M., "The Effects of Factor Prices and Technological Change on the Occupational Demand for Labor: Evidence from Canadian Telecommunications," *Journal of Human Resources*, Spring 1983, *18*, 161–76.

Finch, F. H., "Enrollment Increases and

Changes in the Mental Level of the High-School Population," *American Psychological Association*, June 1946.

Flynn, James R., "The Mean IQ of Americans: Massive Gains 1932 to 1978," *Psychological Bulletin*, 1984, *95*, 1, 29–51.

_____, "Massive IQ Gains in 14 Nations: What IQ Tests Really Measure," *Psychological Bulletin*, 1987, *101*, 2, 171–191.

Forsyth, Robert A., "Achievement Trends for the Iowa Tests of Educational Development in Iowa: 1942–1985," Iowa City: Iowa Testing Programs, personal communication, 1987.

Hauser, Robert M. and Featherman, David L., "Equality of Schooling: Trends and Prospects," *Sociology of Education*, April 1976, *49*, 99–120.

Hieronymus, A. N., Lindquist, E. F. and Hoover, H. D., *Iowa Test of Basic Skills: Manual for School Administrators*, Chicago: Riverside Publishing, 1979.

Husén, Torsten, "The Influence of Schooling Upon IQ," *Theoria*, 1951.

_____, *Talent Opportunity and Career*, Uppsala: Almquist and Wicksell, 1969.

Jencks, Christopher et al., *Who Gets Ahead? The Determinants of Economic Success in America*, New York: Basic Books, 1979.

Johnson, George R., "High School Survey," *Public School Messenger*, 1935, *33*, 3–37.

Jorgenson, Dale W., "The Contribution of Education to U.S. Economic Growth, 1948–1973," in Edwin Dean, ed., *Education and Economic Productivity*, Cambridge, MA: Ballinger, 1984.

_____, **Gollop, Frank and Fraumeni, Barbara**, *Productivity and U.S. Economic Growth*, Cambridge, MA: Harvard University Press, 1987.

Kendrick, John W., "Survey of the Factors Contributing to the Decline in U.S. Productivity Growth," in *The Decline in Productivity Growth*, Sponsored by the Federal Reserve Bank of Boston, June 1980.

Koretz, Daniel, *Trends in Educational Achievement*, Washington: U.S. Congressional Budget Office, 1986.

Lorge, Irving, "Schooling Makes a Difference," *Teachers College Record*, 1945, *46*, 483–92.

_____, "A Table of Percentile Equivalents for Eight Intelligence Tests Frequently Used with Adults," *Journal of Applied Psychology*, 1936, *19*, 392–95.

Nordhaus, William D., "Policy Responses to the Productivity Slowdown," in *The Decline in Productivity Growth*, Sponsored by the Federal Reserve Bank of Boston, June 1980.

Olneck, M., "On the Use of Sibling Data to Estimate the Effects of Family Background, Cognitive Skills, and Schooling: Results from the Kalamazoo Brothers Study," in P. Taubman, ed., *Kinometrics: The Determinants of Socioeconomic Success Within and Between Families*, Amsterdam: North-Holland, 1977.

Roessel, F. P., "Comparative Mental Ability of High School Pupils in Three Minnesota towns in 1920 and 1934," University of Minnesota Press: *Minnesota Studies in Articulation*, 1937, 122–28.

Rundquist, E. A., "Intelligence Test Scores and School Marks of High School Seniors in 1929 and 1934," *School and Society*, 1936, *43*, 301–04.

Swanson, Edward O., "Fashions in Test Scores, or What Goes Up Must Come Down?," in *Student Counseling Bureau Review*, University of Minnesota, September 1973, *25*, No. 1.

Taubman, Paul and Wales, Terence, *Mental Ability and Higher Educational Attainment in the 20th Century*, The Carnegie Commission on Higher Education, NBER Occasional Paper No. 118, 1972.

Tuddenham, R. D., "Soldier Intelligence in World Wars I and II, *American Psychologist*, 1948, *3*, 54–56.

Waters, Brian K., *The Test Score Decline: A Review and Annotated Bibliography*, Technical Memorandum 81-2, Directorate for Accession Policy, Office of the Secretary of Defense, April 1981.

Wheeler, Lester, R., "A Comparative Study of the Intelligence of East Tennessee Mountain Children," *Journal of Educational Psychology*, *33*, 1942, 321–34.

Yerkes, R. M., ed., "Psychological Examining in the United States Army," *Memoirs of the National Academy of Sciences*, *15*, 1921.

U.S. Department of Labor, Bureau of Labor Statistics, "Multifactor Productivity Measures, 1987," Washington: September 30,

1988.

College Board, *The College Board Technical Handbook for the Scholastic Aptitude and Achievement Tests*, Princeton, NJ: College Entrance Examination Board, 1984.

_____, *ATP Guide for High Schools and Colleges*, Princeton, NJ: College Entrance Examination Board, 1987.

U.S. Department of Labor, *General Aptitude Test Battery Manual*, Manpower Administration, 1970.

U.S. Congress, Senate, Committee on the Judiciary, Farr, R. and Tone, B., "What Does Research Show?," *Examination of Basic Skill Achievement Designed to Assist Both States and Local School Districts to Expand and Improve Their Programs in Basic Skills in the Elementary Grades, Testimony before a Subcommittee on Education, Arts, and Humanities of the Committee on Labor and Human Resources*, 96th Cong., 1st sess., Washington: USGPO, 1979.