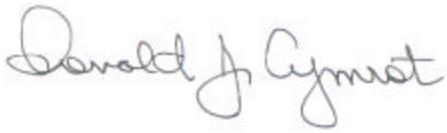
CRM D0008781.A1/Final
August 2003

Does Education Reform Improve Job Performance?

Jennie W. Wenger • Diana S. Lien • Linda C. Cavalluzzo

4825 Mark Center Drive • Alexandria, Virginia 22311-1850

Approved for distribution: August 2003



Donald J. Cymrot, Director
Workforce, Education and Training Team
Resource Analysis Division

This document represents the best opinion of CNA at the time of issue.
It does not necessarily represent the opinion of the Department of the Navy.

Approved for Public Release; Distribution Unlimited. Specific authority: N00014-00-D-0700.
For copies of this document call: CNA Document Control and Distribution Section at 703-824-2123.

Contents

Note: This study was funded by the Department of the Navy, through the general research program at the Center for Naval Analyses (CNA). The CNA Corporation, 4825 Mark Center Drive, Alexandria, VA 22311, is a non-profit research and analysis organization. The opinions expressed in this paper are those of the authors. We are extremely grateful to Cathy Hiatt, Lynda Houck, and Robert Shuford for excellent research assistance. All remaining errors are our own.

Executive summary

The late 1980s and the 1990s were a period of intensive reform within American public education. This period saw the introduction and implementation of a variety of school reforms, including voucher programs, charter schools, financial reforms, increased use of competency tests for both students and teachers, an emphasis on hiring teachers with coursework in the areas they would teach, and increased graduation requirements.

In terms of verified student outcomes, there is little direct evidence that most of these reforms had the desired impact. This is partly because virtually no data exist to allow researchers to tie specific school characteristics to eventual worker quality.[1] Linking K-12 school characteristics to worker outcomes is complicated for a number of reasons. For example, it is difficult to separate the effects of K-12 schooling from those of postsecondary education for students who move directly from high school to postsecondary school.

However, at least four organizations keep quite complete records on performance measures of recent high school graduates. Each year, the four Armed Services enlist roughly 180,000 people. The majority of those who enroll enter the military soon after completing high school. Thus, military records provide a unique, highly detailed source of information on the performance of recent high school graduates. This study focuses on Navy recruits. We match information about recruits' job performance to characteristics of their schools in order to explore the impact of school reform on overall recruit quality, as well as on the success of individual recruits.

Our results suggest that, by most measures, the quality of Navy recruits increased during the 1990s. However, we find that school reform is not responsible for this increase in quality. Specifically, recruits from states that enacted reforms do no better on our outcome measures than recruits from states that did not enact reforms. In fact, in many cases, recruits from "reform" states perform at lower levels than those from states without reforms. Also, reforms seem to affect high school graduates and non-graduates differently; the effect is either less negative or slightly positive for those who graduate with a high school degree, depending on the performance indicator used. These results are unchanged by corrections for unobserved state-level factors. Thus, the increase in quality of Navy recruits appears to be attributable to selective recruiting and/or a general increase in achievement, rather than to any specific school reform(s).

[1] One exception is Card and Krueger (1992); using Census data, these researchers link state-level school characteristics to men's earnings. They find that increased school quality, measured by teachers' education, has a positive impact on both earnings and returns to additional years of education. However, Betts (1995) was unable to find a relationship between school quality and earnings using data on later cohorts.

Background

This section provides background information on education reform and the Navy.

School reforms in the 1980s and 1990s

In April 1983, the National Committee on Excellence in Education published "an open letter to the American people" entitled *A Nation at Risk* (1983). This group, created by the Secretary of Education, set out to quantify the problems within the American primary, and especially secondary, systems, and to come up with specific recommendations. As was clear from the very beginning of the report, the committee focused on America's ability to compete with other countries. The committee's recommendations included increased requirements for high school graduation and college admission, increased requirements and salaries for teachers, and increased monetary support for schools. Many of the reforms of the last 20 years can be traced back to these recommendations.[2]

We classify these reforms into four subcategories. First, this was a period that saw a substantial emphasis on *financial* reform. Until 1973, many education advocates had hoped to challenge disparities in educational spending at the national level, but with *San Antonio Independent School District v. Rodriguez* [411 U.S.1 (1973)] the Supreme Court declared that funding schools with local property taxes did not violate the U.S. Constitution on equal opportunity grounds. At this point, the battle over spending disparities between school districts shifted completely to the state level. In 1971, the California Supreme Court ruled that the California school finance system violated the state's provisions for equal opportunity.[3] The courts interpreted the state Constitution in a variety of ways; between 1971 and 1993, 21 states reformed their education finance systems without court-ordered reform and another 16 states reformed in response to a court order.[4] Evans, Murray and Schwab (1997) discuss the reforms in more detail; one of their findings is that court-ordered financial reform generally results in more equal funding, specifically by raising spending in those districts that previously had low spending. These authors find that reform occurring without a court order is generally ineffective at equalizing funding.[5]

During this same period, there was substantial emphasis on specific reforms within the classroom characterized as involving *increased regulation*. A number of these reforms

[2] This period is not the first era of reform in American schools; Tyack and Cuban (1995) characterize the entire century as one long period of reform and specifically discuss reforms in the 1920s that almost parallel those suggested in *A Nation at Risk.*

[3] *Serrano v. Priest*, 96 California Reporter 601 (1971). Actually, the *Serrano v. Priest* decision of 1971 (known as *Serrano I)* found that California's school finance system violated both state and *federal* provisions; after the *Rodriguez* case, the California Supreme Court reaffirmed its earlier ruling in *Serrano II* and declared the state system illegal based only on the state provisions for equal opportunity.

[4] Evans, Murray and Schwab (1997), p. 11. Many of the states that reformed without a direct order from the courts likely did so to avoid such an order; in two states, however, reform without a court order was later followed by a court-ordered reform.

[5] For a longer history of the evolution of state funding levels and mechanisms for education, see Augenblick, Myers, and Anderson (1997).

were aimed at improving teacher quality; they included more emphasis on professional development, on hiring teachers with substantial coursework in the courses they would teach (especially at the high school level), and on advanced certification for teachers. Along with emphasis on teacher training, a new focus on graduation standards appeared in response to the characterization of American high schools, notably by *A Nation at Risk,* as being lax. School districts and states attempted to increase the "quality" of high school graduates through various competency exams, increases in the number of courses required for graduation, and other means.

Although many saw increased regulation and stringency as the route to improving education, a substantial group believed just the opposite—that only through *decreased regulation* could schools and parents escape bureaucratic tangles to hire the best teachers, select the best curricula, choose the best schools, and increase student achievement. For example, some critics believed that the key to increasing achievement (especially for poor students) lay in giving parents freer choice over which schools their children would attend. Plans to do this included providing school vouchers, using public funds to support quasi-independent charter schools, and establishing various public school choice plans, such as magnet systems. There is little evidence about charter or magnet schools to date, but findings from voucher experiments suggest that results are likely to be mixed (e.g., see Rouse, 1998; Witte et al., 1994) and the charter school movement cannot yet be described as widespread.[6] It is possible, however, that the presence of charter or magnet school laws signaled a commitment to educational reform; therefore, these laws may be associated with increased achievement.

Broadly speaking, the above reforms focus on financial issues or on the level of regulation within schools. One other type of reform occurred during the same period. In an influential set of papers, Smith and O'Day (1991, 1993) postulate that narrowly targeted reforms (such as those aimed at teacher or student coursework) will have little impact on overall school quality because these reforms cannot fundamentally change the way schools function. Rather, these authors advocate a type of bottom-up, or total, reform that focuses on changing the way schools teach, and thus the way students learn. This is often referred to as *systemic reform*. This type of reform includes some of the ideas advocated by the previously mentioned groups, such as improved professional development (teacher training), and increased student access to certain key math and science classes. A central focus of systemic reform is the alignment of both teacher training and classroom instruction with appropriate instructional materials and accepted standards for learning. At the same time, systemic reform recognizes the importance of involving parents and community leaders in successful school reform. The model also postulates that the synergy between the different components is of fundamental importance to successful reform. Because of NSF funding, systemic reform was instituted in a number of states during the 1990s. The programs were relatively small in scope, but, through the focused effort of one influential organization, systemic reform may have

[6] Even Arizona, the state with the highest number and highest density of charter schools, had only one school for every 3,300 students (the median enrollment in charter schools in 1998-99 was 132 students (U.S. Charter Schools, 2000). States with the lowest number of schools per 10,000 students include Mississippi with one charter school and New York with five schools.

influenced the quality of education in the states where it was instituted.[7] Few studies link systemic reform to student outcomes, but Grissmer et al. (2000) suggest that systemic reform may be responsible for increased standardized test scores in some states.

Thus, in general terms, four different types of reforms occurred during the 1990s. We define each type of school reform for our analysis as follows: A state had *court-ordered financial reform* if the ruling from a legal case required the state to change its funding system on equity grounds.[8] We measure increased regulation by requiring that students *pass a test to graduate.* We define decreased regulation by the existence of a law permitting *charter schools*. We define *systemic reform* by the existence of a grant from NSF under its State Systemic Initiative (SSI) program to institute systemic reform.[9] Table 1 indicates the number of states in which each type of reform occurred.

Table 1. School reforms enacted as of 1998

Type of reform	Number of states with reform by 1998
Financial (change in manner of funding districts):	
Court-ordered [a]	16
Not court-ordered [a]	21[d]
Increased regulation:	
Graduation standards—must pass test to graduate [b]	17
Decreased regulation:	
Charter school legislation in place [c]	36
Systemic reform:	
State had systemic program	25

Notes/Sources:
a—Evans, Murray, and Schwab (1997).
b—Digest of Education Statistics (1999), table 158 (figures for 1995-6).
c—U.S. Charter Schools (2000).
d—Two states had both non-court-ordered and (later) court-ordered reforms; by 1999, 15 states had neither type of reform in place [Evans, Murray, and Schwab (1997)].

The 1980s and 1990s can be classified as a period of intense, diverse school reforms. Students who completed school in the late 1980s were probably affected only minimally by these reforms; changes in response to *A Nation at Risk* began to happen in the middle to late 1980s, and charter schools and systemic reforms began in the 1990s. In contrast,

[7] Yearly systemic reform funding was about $100 million; the awards were spread between states, cities, and districts. There are roughly 90,000 public schools in the United States, so on average each school could expect to receive about $1,100. Overall spending on public elementary and secondary schools in the United States is over $350 billion.
[8] A state had *non-court-ordered financial reform* if that state changed its funding system to increase equity, without a court order. Note that all non-court-ordered reforms occurred before 1989, the first year of our Navy data.
[9] As of 1994, 25 states had such programs; by 1999, many of the grants were over, and only 8 states had a program (like NSF, we expect the effects of systemic reform to last beyond the grant period).

these reforms are most likely to affect students who completed high school in the late 1990s; these students entered school late enough that many of the financial and other reforms were in place. Therefore, we measure the impact of these reforms by comparing students who completed school in the late 1980s with those who did so in the late 1990s.

Standardized tests

Everyone who wants to enlist in the Armed Forces take the ASVAB (Armed Services Vocational Aptitude Battery) exam. This multiple-choice exam is comprised of ten subject-specific tests. Scores from the tests measuring vocabulary and math skills are combined to yield an AFQT (Armed Forces Qualification Test) score. This score is important because those with a high school diploma must score at or above the 30th percentile to be admitted to the Navy. The requirement for those without a high school degree is higher; they must score at or above the 50th percentile. In addition, minimum scores on certain subtests are required to qualify for certain Navy occupational fields (ratings). In particular, two subtests measure knowledge in science and mathematics; we examine them in some detail in a later section.[10] Thus, the distribution of scores on the AFQT portion of the ASVAB exam provides some information about the quality of high school graduates.[11]

An important drawback to the AFQT, however, is that the sample of those who take the test is not random. Thus, changes in average performance cannot be generalized to the full population from which the sample is drawn without taking into account the selection process.[12] However, comparing AFQT scores over this period to scores on a standardized test given to all students provides information about the changing (relative) quality of military recruits. The most appropriate comparison is the National Assessment of Educational Progress (NAEP) test, which is periodically given to a random group of students in each state. The NAEP assessments consist of two different types of tests. The Trend Assessment administers the same tests with the same questions to students to provide an absolute baseline; the tests began in 1969. In general, the trend is flat over time. The Main National NAEP exams change to reflect changes in the curriculum over time. These test scores show slight upward trends over time.[13]

---

[10] We norm the subscores to allow direct comparison across years.

[11] The subject of exactly *what* the AFQT measures is the focus of a small body of research. Some suggest that it measures intelligence (e.g., Herrnstein and Murray, 1994); a strong argument against this position is provided by Gould (1981). Some researchers cite the exam's potential bias (e.g., Rodgers and Spriggs, 1996). Others use it as a measure of job skills (Neal and Johnson (1996) refer to it as a measure of "acquired skill"). Although some research suggests the test does a good job at predicting on-the-job performance, Cooke and Quester (1988) suggest that high school completion is probably a more important predictor of productivity. Nonetheless, a person's AFQT score is widely considered to be positively correlated with what a student learns in school, and highly correlated with success in the military.

[12] The sample of those taking the AFQT tests includes those who have expressed an interest in the military, plus students in certain schools that elect to give the test to all enrolled students.

[13] The 8th grade Trend NAEP for mathematics emphasizes arithmetic and basic skills, whereas the Main National exam emphasizes problem solving and geometry. For more information, see: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000469

A second comparison can be made between AFQT and SAT scores to explore how the relative sizes and quality of those bound for postsecondary school versus the military have changed over time. Overall, SAT scores fell through the 1970s and the early part of the 1980s before leveling off. Scores increased slightly during the middle to late 1990s. Most or all of the initial decrease, however, can be explained by an increase in the number/proportion of students taking the test. There was virtually no change in scores of those from the highest quintile during the 1970s and 1980s; in fact, the verbal scores trend upward slightly beginning around 1985.[14] Therefore, it seems reasonable to say that, as measured by SAT scores, average achievement held steady or increased slightly over the time period in question.

Combining information from the SAT and NAEP exams suggests that overall achievement was fairly constant between 1989 and 1999, with possible small improvements near the end of the time period. Thus, all else equal, the pool of potential Navy recruits may have improved slightly over time. Along with this trend, of course, the Navy recruited fewer Sailors in 1999 than in 1989, so it seems probable that the average quality of a Navy recruit would have improved slightly over the period without school reform or changes in the civilian economy.

The Navy in the 1980s and 1990s

Recruit quality is of great concern to the Navy; substantial evidence suggests that "high-quality" recruits (those with a high school diploma and a score of over 50 on the AFQT) exhibit significantly lower levels of attrition than other recruits.[15] The Navy also makes concerted efforts to limit the number of Sailors who do not fulfill their initial obligation (those who "attrite"). Attrition is extremely expensive to the Navy. Estimates indicate that recruiting alone costs roughly $10,000 per recruit; Sailors who leave without fulfilling their obligation also receive costly training, for which the Navy receives little or no return on its investment.

The number of Navy recruits varies somewhat from year to year; in general, the number was higher in the late 1980s and early 1990s than today. The number of recruits then fell sharply during the military "drawdown" from nearly 95,000 in 1986 to less than 50,000 in 1996. The number of recruits has remained stable at roughly 50,000 over most of the last 5 years. The drawdown brought large decreases in the number of accessions, and small increases in average AFQT scores. Our challenge is to separate the effects of fewer accessions from those of school reform and those of changing civilian opportunities. (Because the drawdown represents a period when expectations were dramatically different, we exclude this period from our individual-level analysis, instead focusing on the differences between 1989 accessions and 1999 accessions.)

[14] Data on state-level average SAT scores and the proportion of students taking the exam are available in the *Digest of Education Statistics* (various years). See Berliner and Biddle (1995) for a detailed analysis of SAT scores by high school class rank.
[15] See, for example, Garcia et al. (2001).

Sample selection

Several types of sample selection could affect our results. First, certain states, probably those with relatively low-quality schools, are especially likely to institute reforms. Second, those who enlist in the Navy are not a random subsample of American high school students. This means it may not be appropriate to generalize our results to all high school students. Also, this could pose a particular problem for our research if the types of students who enlist in the Navy change over time. We might expect this to be the case if, for example, civilian opportunities changed over this time period. For this reason, we control for civilian labor market conditions and for postsecondary opportunities.

First, we determine what types of state were most likely to enact reforms. Although some states are clearly more committed to education reform than others, there is no simple correlation between all these types of reforms. In other words, states that have financial reform are not more likely to have charter legislation than those without financial reform. In general, states with court-ordered or non-court-ordered financial reforms look very much like states without such reforms. States requiring that students pass a test to graduate have lower SAT scores, and more students taking the SATs, than other states. This is also true in states with charter legislation, as well as those that ever had an SSI program in place. In addition, states with SSI programs tend to have higher incomes and poverty levels (more income disparity) than states without these programs. Some of these differences are regional in nature; students in the South and the Northeast are more likely to take the SAT than other students because (a) colleges they are likely to attend require it and (b) both southern and northeastern states are overrepresented among states with an SSI program. Overall, the data suggest that there may be some negative sample selection among the reform states. In the analyses that follow, we use state fixed-effects models to control for baseline differences in student outcomes across states that may be correlated with enactment of reforms.

Next, in table 2, we examine changes in civilian opportunities during this time frame. One might view high school seniors as choosing between three options: enlist in the Armed Forces, attend postsecondary school, or join the civilian labor force. Statistics suggest that, during the 1990s, opportunities in the civilian labor force and in postsecondary institutions improved. First, real (inflation-adjusted) monthly wages increased in both the manufacturing and the service sectors (see table 2). The national unemployment rate fell from 5.3 in 1989 to 4.2 in 1999, as did the unemployment rate for 16- to 19-year-olds (from 15 to 13.9 percent). At the same time, college enrollments increased, and the percentage of 18- to 24-year-olds enrolled in degree-granting institutions increased by 15 percent. Therefore, all else equal, we might expect the quality of Navy recruits to have fallen during this period. However, the number of recruits also fell; to some extent this should mitigate the effects of improving civilian opportunities. Moreover, military pay increased during this period; after adjusting for inflation, basic pay increased nearly 13 percent for an E-3 with 2 years of service. Basic pay does not represent total compensation that Sailors receive. For example, depending on marital status, Sailors receive different levels of housing allowances. and the like. Military increases in compensation, however, typically come in the form of basic pay raises. Therefore, the 1990s represent a period of improved opportunities both within and

outside the military, as well as a period in which the Navy's total demand for Sailors decreased.

Table 2. Military and civilian opportunities during the 1990s

	1989	1994	1999
Basic Military Pay (real monthly, E-3 with 2 years of service)	$1,046	$1,113	$1,179
Average Real Monthly Pay, Manufacturing Sector	$1,181	$1,195	$1,256
Average Real Monthly Pay, Service Sector	$840	$846	$944
Average Civilian Unemployment Rate	5.3%	6.1%	4.2%
Average Unemployment Rate, 16- to 19-year-olds	15.0%	17.6%	13.9%
Percentage in Postsecondary School	30.9%	34.6%	35.6%

Successful education reform could affect the quality of Navy recruits in at least three ways. First, if all high school students are (or the average high school student is) more competent or better prepared as a result of reform, all else equal, this would cause the average Sailor to be more competent and productive. However, successful school reform could also change opportunities and aspirations of would-be recruits, resulting in no impact or even a negative impact on the quality of new Sailors. For example, systemic reform could promote college attendance among high school students. If school reform increases the number of students who achieve at the level necessary to attend college (i.e., school reform shifts the achievement of all students), this could leave the Navy with recruits of the same quality as before, although these would be different people than the Navy would recruit before. For example, we might expect more minority Sailors and Sailors from disadvantaged backgrounds under this scenario. Another possibility is that school reform could not only increase achievement but also increase motivation and access to higher education—students at an achievement level who previously would not have applied to college may apply now and attend. In this case, the quality of new Sailors could be expected to be lower than before the (successful) school reform. Finally, it is possible that these countereffects are small in size, causing education reform to have little or no effect on the quality of new recruits.

Because of the changes occurring within the civilian economy and the Navy during the 1990s, as well as the role that reform may play in influencing career choices, examining the success of the individual Sailor cannot tell us the whole story about how school

reform affects Sailor quality. To get at these issues, we would like to model the choice that high school seniors make among civilian employment, college, and the military. Unfortunately, our data are limited to ASVAB test-takers and to Navy recruits. Using these data, our approach instead is to first examine data on all who took the ASVAB exam. Although tastes for military service probably changed over this period, these data give us a relatively clean measure of how overall achievement changed. Next, we use state-level models to examine the relationship between various types of reform and quality of Sailors. Finally, we use individual-level data to see how attending school in a state where reforms were instituted affects the probability that a Sailor will succeed in the Navy. Although this approach does not fully capture the mechanism by which reform affects recruits, it does allow us to control for factors that influence the selection process and to quantify the impact of reform on military test-takers and Navy recruits.

We use the data on individual recruits to measure several different outcomes. To focus on recruit quality, we measure the effect of school reforms on AFQT scores. In addition, we examine how school reforms affect the probability that an individual Sailor will leave the Navy before fulfilling his/her obligation. We also examine how Sailors perform on the job. In particular, Sailors must fulfill certain additional requirements in order to be rated (placed in a specific job); we examine the effect of school reforms on the probability (a) that a Sailor is rated and (b) that a Sailor qualifies for a technical rating within 23 months of active duty. Finally, we look at promotion, examining how school reforms affect the probability that a Sailor will achieve a given paygrade within the first years in the Navy.

Data and descriptive statistics

We formed our dataset from Defense Manpower Data Center (DMDC) extracts of all who took the ASVAB during fiscal years 1988, 1993, and 1998. (These are the people who could have joined the Navy during 1989, 1994, and 1999, respectively. Throughout this section, we refer to the subsamples by the year in which they joined the Navy). We use the set of all test-takers, as well as all test-takers who state an interest in the Navy, in descriptive statistics below. We excluded some from this group because of missing information.[16]

In 1994 and 1999, the number of ASVAB test-takers dropped dramatically in proportion to the number of test-takers in 1989 and in proportion to the total population aged 18 to 24.[17] The decrease in the number of test-takers does not seem to have been a response to the strength of the civilian economy. In 1994, the civilian economy was weaker, and proportionally more test-takers were from states with higher unemployment rates. (In 1994, the average state unemployment rate was 6.1 percent.) AFQT test scores peaked in

[16] We excluded those who did not have a home state listed or had a home state of the District of Columbia or a U.S. territory. In addition, we excluded those who are missing information on their education level, or gender. We also excluded test-takers younger than 16 years of age and older than 21 years of age. For all ASVAB test-takers and Navy Propensity test-takers, the demographic information is from the DMDC ASVAB extract.

[17] In 1989, there were roughly 13.2 million youth between the ages of 18 and 24 in the United States; by 1999, that figure had increased to 13.7 million. Therefore, a much smaller proportion of people elected to take the ASVAB exam in 1999 than in 1989.

1994. The percentage of test-takers who were high school dropouts also decreased during this period. These findings, shown in table 3, suggest that the mid-1990s military drawdown made FY 1994 a year that included particularly less military recruiting efforts and stricter recruitment eligibility rules.

Table 3. Descriptive statistics, all ASVAB test-takers

Variable	1989	1994	1999
Average AFQT score[a]	51.1	55.0	51.1
Percentage scoring below 30	20.0	12.7	17.4
Percentage scoring at least 80	13.3	15.5	11.8
Caucasian	69.7	71.5	65.7
Hispanic	6.2	8.5	10.4
Male	83.6	79.5	77.7
Some college	1.4	1.7	2.8
High school dropout	6.9	1.7	1.3
State unemployment rate	5.5	6.1	4.3
Percentage of teachers with MA	39.9	41.1	41.6
Charter legislation	0.0	25.6	83.9
Court-ordered financial reform	29.0	38.0	40.3
Required to pass test to graduate	0.0	0.0	50.1
Number of observations	414,845	194,515	260,936

[a]AFQT scores are constructed from a subset of the 10-subtest ASVAB. We use the same subtest composition to compare scores over time. The AFQT composite scores are reported as percentiles and are based on a 1980 reference population.

Table 4 shows that those test-takers who expressed an interest in the Navy are similar to the entire test-taker sample in terms of overall average test scores. Consequently, we focus on all test-takers and those test-takers that joined the Navy. Our first regression results aggregate data to the state level for all who took the AFQT during FY 1988, FY 1993, or FY 1998. We then examine state-level data for all who enlisted in the Navy within 2 years of taking the AFQT. We exclude some from this group because of various missing information.[18] Also, for our individual-level analysis, we exclude those who took the exam during FY 1993 because of concerns that the atmosphere in the Navy was different during this period in unmeasured ways. After these exclusions, we are left with roughly 92,000 observations. We present some descriptive statistics on the sample of people who joined the Navy in table 5.

[18] Specifically, we excluded those who had no home state listed as well as those who were educated in the District of Columbia and in U.S. territories. We also excluded those who were missing information on their education level or length of service, and those who were either less than 18 years of age or more than 21 years of age at accession. This last exclusion serves to increase the accuracy of the educational information we assign to recruits; men and women are allowed to join the Navy up to age 35, but these people obviously attended school at a different time period than the majority of the sample. We have two sources of information for some variables—specifically, those measuring gender, race/ethnicity, and education. In these cases, we use information from CNAC's Enlisted Master Record (EMR). In cases when EMR information is missing, we instead use information included in our DMDC data extract.

Table 4. Average test scores

	All ASVAB Test-takers [a]	Navy Propensity Test-takers [a]
Average AFQT score	52.0	51.9
Percentage scoring below 30	17.6	18.0
Percentage scoring at least 80	13.3	14.0
Number of observations	870,296	228,844

[a] Includes 1989, 1994, and 1999 data

Table 5. Descriptive statistics, sample of Navy accessions (individual level)

Variable	1989	1999
Average AFQT score	53.8	57.0
Percentage scoring below 30	14.2	1.8
Percentage scoring at least 80	15.0	13.7
General Science ASVAB[a]	51.3	51.1
Math Knowledge ASVAB[a]	52.1	54.4
Caucasian	74.8	66.0
Black	21.8	20.3
Hispanic	8.7	11.2
Male	88.2	81.4
Some college	.88	2.7
High school dropout	5.1	4.6
State unemployment rate	5.5	4.3
Western state	20.3	27.2
Percentage of teachers with MA	39.5	40.1
Charter legislation	0	85.3
Court-ordered financial reform	29.5	43.0
Required to pass test to graduate	0	48.0
Boot camp attrition	9.2	15.9
12-month attrition[b]	9.4	9.8
23-month attrition [c, d]	7.0	7.1
Proportion rated[d]	57.1	57.7
Proportion rated in tech rating[d]	35.8	33.9
Promoted to E3 in 12 months[c]	21.8	43.6
Promoted to E4 in 23 months[e]	24.9	29.2
Number of observations	59643	32221

Notes on table 5: a—ASVAB subtest scores were converted to a common scale based on the 1980 youth population. These standardized scores have a mean of 50 and standard deviation of 10 for that population. b—Calculated conditional on remaining in the Navy through boot camp. c—Calculated conditional on remaining in the Navy for 12 months. d—Calculated for all who joined the Navy at least 23 months before September 2001. e—Calculated for all who joined the Navy at least 23 months before September 2001, conditional on remaining in the Navy at least 23 months.

As table 5 reveals, the demographics of new Sailors changed markedly over this period. By 1999, just over half of all new Sailors were women and/or minorities (defined as African American, Asian/Pacific Islander, American Indian, Hispanic, or Other, non-white). At the same time, the average AFQT score increased somewhat over this period because the Navy recruited far fewer people with very low AFQT scores in 1999 compared with 1989. The overall civilian economy was stronger in 1999 than in 1989, as evidenced by the average state-level unemployment rates. Also, there were regional shifts in the home states of recruits over this period; most noticeably, by 1999 27 percent of recruits came from western states as opposed to 20 percent in 1989 (the proportion of recruits from northeastern and midwestern states fell over the time period).

In terms of Navy outcomes, 1999 was a year of historically high attrition, especially during boot camp. Beyond boot camp, attrition rates were similar between 1989 and 1999, and roughly the same proportion of Sailors was rated within 23 months in both years. One other change between 1989 and 1999 concerns promotion—by 1999, a Sailor was much more likely to reach the E-3 paygrade within 12 months than in 1989, but the rates of promotion to E-4 were similar during the 2 years.[19]

Results

In the first phase of our analysis, we examine state-level AFQT scores for people who took the ASVAB (test-takers), and for those who took the ASVAB and joined the Navy (accessions). For each group, we examine the effects of individual- and state-level characteristics on AFQT scores. In the second phase of our analysis, we examine individual-level data on accessions. Again, we look at the effects of both individual- and state-level characteristics on individual AFQT scores. Comparing and contrasting the state-level and individual-level results provides information about how education reform (and other factors) affects the distribution of test-takers and Sailors. Finally, we complete the second phase of our analysis by examining how education reform affects other job-related outcomes, such as attrition and promotion. All of these regressions are production functions; we assume that individual- and state-level characteristics combine to "produce" test scores and other outcomes. This approach is common in education literature; see Hanushek (1989) for a review of the numerous education production function studies.

State-level results

For 1989, 1994, and 1999, we estimate whether state-level reforms influenced the overall quality of all ASVAB test-takers and Navy accessions. All state-level results include several classes of control variables:

[19] A person at the E-3 paygrade in the Navy is referred to as a Seaman; this is equivalent to a Private First Class in the Army. A person at the E-4 paygrade is referred to as a Petty Officer Third Class, which is equivalent to an Army Corporal.

- State-averaged individual demographic controls: gender, race, ethnicity, and high school graduation percentage (see tables 3 and 5 for detailed descriptions of these variables).[20]
- State-level indicators of alternative opportunities and demographic controls: poverty rate, unemployment rate, proportion in a metro area, in-state college tuition, number of seats in public postsecondary schools per 100 graduating high school seniors, and number of test-takers.
- Education controls: per-pupil spending, proportion of spending from local and state sources, teacher education and experience levels, and pupil-teacher ratio.
- Education reform: law permitting charter schools, court-ordered financial reform, non-court-ordered financial reform, test to graduate, and state systemic reform initiative.[21]
- Additionally, as noted, specifications include year and state dummy variables.[22]

As discussed above, those who join the Navy are a nonrandom group. For example, on the demand side, the Navy has accession restrictions based on education and AFQT score. Thus, the impact of educational reforms may affect Navy accessions differently than the average student. Though we don't have outcome information on all students, we can examine the impact of school reform on the broader sample of all AFQT test-takers.

In our naïve model of all test-takers, we estimate the impact that state-level educational reform had on average AFQT scores, without state- or year-level dummies (see table 6, column 1). For the naïve model, we find that states with charter legislation had on average a 1.5-point lower AFQT score than other states. States with systemic reform had a .6-point lower score, and states with court-ordered financial reform had a .8-point lower score. This could imply that the implementation of reform reduced AFQT scores, or that states with systemic, charter legislation, or court-ordered reform had, on average, lower AFQT scores than other states. To account for permanent state-specific characteristics and to estimate the impact of educational reform within a state, we next include state dummies (column 2). We find that the estimated effects of charter legislation, court-based finance, and systemic reform were the result of these reform states differing from other states. Thus, when we account for time-invariant state characteristics, we find no impact of reform. Finally, we add year fixed effects; again, reforms have only small, insignificant effects on AFQT scores (see column 3).[23]

Navy accessions are not representative of all AFQT test-takers: they have higher AFQT scores on average and had a larger percentage increase in AFQT scores over this period than all test-takers. Because Navy accessions are of higher "quality," they may be affected by educational reforms differently than all test-takers. For example, as a group

[20] With all test-takers, we included a control variable of the state-year proportion of test-takers who expressed an interest in the Navy over the three other military services.
[21] States that implemented non-court-based financial reform did so before 1989. Because there is no variation in this variable, we do not include it in the state and year fixed-effects models.
[22] Along with controlling for the log of state-year cell observations per 18- to 24-year-olds, we weighted our observations by the number of observations by state-year.
[23] We find the same result for the subsample of test-takers that expressed an interest in the Navy over the other services, but didn't all join the Navy.

they may be more likely to perceive postsecondary education, as opposed to getting a civilian job, as their next best alternative to joining the Navy. If reforms are effective, they may increase the overall quality of accessions or decrease high-quality students' interest in the Navy. Because Navy accessions are the population of interest, we next examine how education reform affected AFQT scores of this group.

Table 6. State-level results—Did reforms affect state-level AFQT scores?

All Test-takers		**AFQT Scores**	
Model→	(1)	(2)	(3)
Test to Graduate	-0.54	0.17	0.30
	(0.42)	(0.38)	(0.31)
Charter Legislation	-1.50**	-0.51	-0.21
	(0.33)	(0.26)	(0.20)
Court-Ordered Financial Reform	-0.78**	0.05	0.52
	(0.39)	(0.61)	(0.53)
Non_Court-Ordered Financial Reform	0.28		
	(0.40)		
Systemic Reform	-0.59*	-0.28	-0.33
	(0.35)	(0.25)	(0.22)
F-statistic[a]	9.7	2.07	2.19
	[0.00]	[0.09]	[0.08]
State Fixed Effects	N	Y	Y
Year Fixed Effects	N	N	Y

Unless noted otherwise, standard errors are in parentheses.
* 10% level of significance ** 5% level of significance.
[a] F-statistic [p-value] on the joint significance of the reform variables.

Table 7 shows our state-level results for Navy accessions. The preferred specification, shown in column 3, indicates that most reforms have no effect on state-level AFQT scores of accessions. Specifically, we find an imprecise positive increase in state-level test scores from charter legislation and no impact of financial reform or systemic reform. Our only precise estimate is from test-to-graduate reform, which decreased average AFQT scores by .8 point.

To examine whether reforms may have had more of an impact on specific subject areas, we estimated the impact of education reforms on state-level General Science and Math

Knowledge ASVAB scores for Navy recruits. Systemic reform focused on these key subject areas, and the Navy puts a premium on recruits' technological knowledge. In this regression, test to graduate again provides the only precise estimate; this reform is predicted to decrease General Science scores by 0.3 point. The general negative result from test to graduate seems counterintuitive because the goal of the high school graduation competency exam was to increase the quality of graduates. However, states implemented test-to-graduate reform after the mid-1990s military drawdown era. When we rerun the model with only 1989 and 1999 years, we find no effect on test scores from test-to-graduate reform (not shown). Thus, it is unclear whether test to graduate decreased the quality of recruits, or whether the number and quality of recruits during the military drawdown differed by these states. Overall, we find only very small, insignificant effects of educational reform on student quality. Also, the educational reforms taken as a group do not influence state-level average scores.[24] Next, we turn to our individual-level results.

Individual-level results

All individual-level results include several classes of control variables:

- Individual demographic controls: gender, race, age, education level.
- State-level demographic controls: poverty rate, unemployment rate, proportion of population with college education, proportion in a metro area, region, number of test-takers.
- Education controls: per-pupil spending, proportion of spending from local and state sources, teacher education and experience levels, pupil-teacher ratio.
- Education reform: law permitting charter schools, court-ordered financial reform, test to graduate, and systemic reform initiative.
- In addition, except where noted, specifications include year and state dummy variables.
- All regressions have Huber-White corrected standard errors, corrected for clustering at the state level.

Upon enlisting, recruits are divided into tiers based on educational attainment. Tier I recruits are recruits with a traditional high school degree or an equivalent certification. Tier II recruits have some form of alternate certification, such as General Educational Development (GED) credentials; Tier III recruits have not completed high school and hold neither a degree nor any alternate certification. We run all regressions separately for Tier I versus Tier II-III recruits.

[24] When accounting for state and year effects, we are unable to reject that the reforms are jointly equal to zero at a 5-percent level of significance. We find the same result for all test-takers and for just those test-takers who expressed an interest in the Navy.

Table 7. State-level results—Are differences in state-level recruit quality a result of educational reform?

Navy Accessions	AFQT Scores			General Science Subscore			Math Knowledge Subscore		
Model→	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Test to Graduate	-0.18	-1.19**	-0.79*	-0.37*	-.52**	-0.34*	0.30*	-0.09	-0.08
	(0.52)	(0.43)	(0.46)	(0.20)	(0.23)	(0.18)	(0.16)	(0.20)	(0.21)
Charter Legislation	0.30	0.75*	0.64	-0.04	0.18	0.19	0.26	0.30	0.15
	(0.47)	(0.41)	(0.41)	(0.18)	(0.16)	(0.15)	(0.19)	(0.20)	(0.17)
Court-Ordered	-0.53	0.97	0.37	-0.50*	0.40	0.27	-0.06	0.62**	0.28
Financial Reform	(0.62)	(0.73)	(0.72)	(0.19)	(0.45)	(0.31)	(0.25)	(0.30)	(0.28)
Non-Court-Ordered	-0.16			-0.09			-0.04		
Financial Reform	(0.50)			(0.16)			(0.20)		
Systemic Reform	0.36	-0.38	-0.44	0.04	-0.05	-0.08	0.14	-0.21	-0.30
	(0.49)	(0.39)	(0.36)	(0.17)	(0.20)	(0.14)	(0.19)	(0.18)	(0.16)
F-statistic[a]	0.49	2.52	1.22	2.90	1.64	1.22	1.61	2.06	0.68
	[0.78]	[0.05]	[0.31]	[0.03]	[0.17]	[0.31]	[0.18]	[0.09]	[0.61]
State Fixed Effects	N	Y	Y	N	Y	Y	N	Y	Y
Year Fixed Effects	N	N	N	N	N	Y	N	N	Y

Unless noted otherwise, standard errors are in parentheses. * 10% level of significance ** 5% level of significance.

[a]F-statistic [p-value] on the joint significance of the reform variables.

AFQT scores

Initially, we present a "naïve" specification in which we do not control for state-specific effects. In this specification, shown in table 8, each type of educational reform has a negative effect on the person's AFQT score. The coefficients indicate that Tier I recruits in states without reforms will have AFQT scores that are 1.2 points *higher* than those in states with all four types of reforms, all else equal. (The effect is similar for Tier II-III recruits; their scores are predicted to be 1.8 points higher without reforms). In the case of Tier I recruits, however, only court-ordered reform is significant at the 5-percent level; in the case of Tier II-III, only test to graduate is significant. For Tier I recruits, the four types of reform are jointly significant at the 6-percent level, but the four variables are not jointly significant for Tier II-III. Therefore, in the case of Tier I recruits, we can say within this model that reforms are jointly associated with lowered test scores; in the case of Tier II-III, we can say only that the presence of a test to graduate is associated with lower AFQT scores (the size of this effect is about 1.2 points, or about one-tenth of a standard deviation). Thus, overall effects are small, especially for Tier II-III recruits. (For all further specifications, we discuss individual and joint effects of reforms only when the effects are significant.)

Table 8. Individual results—How educational reforms affect recruits' standardized test scores

	AFQT Scores (Naïve model)		AFQT Scores		Math Knowledge Subscore	
Sample→	Tier I	Tier II-III	Tier I	Tier II-III	Tier I	Tier II-III
Test to Graduate	-0.24	-1.24**	-0.65*	-2.05**	-0.21	-0.09
	(0.27)	(0.54)	(0.35)	(0.98)	(0.13)	(0.41)
Charter Legislation	-0.28	-0.25	1.19**	0.02	0.75**	-0.02
	(0.38)	(0.76)	(0.46)	(1.28)	(0.19)	(0.50)
Court-Ordered	-0.55**	-0.23	1.51**	-0.93	1.53**	-1.10*
Financial Reform	(0.21)	(0.44)	(0.67)	(1.80)	(0.29)	(0.66)
Systemic Reform	-0.12	-0.08	-0.53	0.21	-0.30	-0.32
	(0.31)	(0.66)	(0.45)	(0.94)	(0.20)	(0.42)
F-statistic[a]	2.29	1.49	2.19	3.13	9.45	1.14
	[0.06]	[0.20]	[0.08]	[0.02]	[0.00]	[0.35]
State Fixed Effects	N	N	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y

Unless noted otherwise, standard errors are in parentheses.
* 10% level of significance ** 5% level of significance.
[a] F-statistic [p-value] on the joint significance of the reform variables.

This naïve specification, however, does not account for unchanging state-specific effects. For example, if states with lower test scores were more likely to face court-ordered reform, even if these states improved following reform, they may still perform below states that did not face court-ordered reform. Similarly, the perceived value of public education may vary quite a bit between states. If, as seems likely, the perceived value affects the probability that the state will enact reforms, the coefficients from the naïve specification may be biased. Therefore, we next estimate an alternate specification in which we allow for state-level fixed effects. These results, also shown in table 8, indicate somewhat different reform effects. In this specification, the effects of charter legislation and court-ordered reform on Tier 1 recruits are positive; together, these reforms are predicted to raise test scores nearly 3 points. In contrast, for Tier II-III recruits, test to graduate alone is predicted to lower scores more than 2 points. This result runs counter to our expectations and may be capturing other factors that we have not accounted for in the model. The other reforms have mixed signs for this group; enacting all four is predicted to decrease scores by roughly 2.8 points (about one-fifth of a standard deviation) overall. Therefore, reforms seem to have had a small positive effect on the test scores of high school graduates, while having a small negative effect on the scores of those who did not complete high school.

Reforms also seem to have some effect on at least one ASVAB subtest. Although reforms have little effect on scores on the General Science or Arithmetic Reasoning subtests for Tier I or Tier II-III recruits, reforms do have an effect on the Math Knowledge subtest score.[25] In particular, the presence of charter legislation and court-ordered reform together are predicted to raise scores on this subtest for Tier I recruits by about 2.3 points (more than .25 standard deviation). For this group, the effects of test to graduate and systemic reform are estimated to be negative and insignificant, so the effect of enacting all four reforms is estimated to raise scores by about 1.8 points. For Tier II-III recruits, court-ordered reform is predicted to lower scores slightly.

Many of the other coefficients in these regressions have the expected signs. (Please see appendix B). Focusing on the AFQT regressions, 18-year-olds generally do better on the test than those who take it later. Among Tier I recruits, those with some college credit generally do worse on the test. This is because many recruits with some college credit actually do not have high school diplomas; college credit is accepted in lieu of a high school diploma. Those with an alternate certificate also score lower than "traditional" high school graduates. Recruits who are members of racial/ethnic minorities score lower than non-Hispanic whites on average; this is consistent with prior research on the AFQT (Jencks and Phillips, 1998). As expected, those who come from states where a higher proportion take the AFQT generally have lower scores (Berger and Toma (1994) demonstrate a similar result for the SAT). The coefficient on per-pupil expenditure is small and insignificant, but the source of funding has an important impact—those from states with more local or more state financing generally achieve higher scores.[26] The effects of teachers' education and experience are small and insignificant, but

[25] The exception for Tier II-III recruits is that test to graduate has a small, negative effect on science scores.
[26] In general, all expenditures come from state, local, or federal sources. The federal component is small, making up less than 10 percent of spending in most cases, but states with more disadvantaged populations usually receive more federal funding.

pupil-teacher ratios have a negative, significant effect on AFQT scores for recruits in all tiers. These results indicate that lowering the average class size from 16 students to 13.8 students (a decrease of one standard deviation from the mean class size in 1999) is predicted to increase average AFQT scores by roughly three-quarters of one percentage point but would require about 500,000 additional teachers.[27] Thus, the effect of pupil-teacher ratios on AFQT scores is small.

The coefficient on the year dummy variable indicates that AFQT (and subtest) scores for Tier II-III recruits were higher in 1999 than in 1989, after correcting for demographic and school characteristics. This suggests that Navy recruits, like students as a group, had slightly higher achievement in 1999 than in 1989. However, this increase is not associated with any of the education reforms examined here.

Attrition

In general, educational reforms have small effects on attrition rates for Tier I recruits (see table 9). We examine three different attrition measures—boot camp attrition, 12-month attrition (conditional on remaining in the Navy through boot camp), and 23-month attrition (conditional on remaining in the Navy for 12 months). It is most likely that school reform would affect 12- and 23-month attrition because recruits attend classroom training after boot camp; therefore, at least some post-boot-camp attrition is academic in nature. Consistent with this, we find that school reform has no effect on boot camp attrition for Tier I recruits; however, reform is associated with increased boot camp attrition for Tier II-III recruits. This effect is driven by charter legislation and test to graduate; the presence of these two types of reform is predicted to substantially increase boot camp attrition.[28]

Generally speaking, school reform is associated with higher 12-month attrition for all recruits. Enacting all four reforms is predicted to increase attrition levels by about 3 points for Tier I recruits, and by about 12 points for Tier II-III recruits. (In an exception to the general pattern, systemic reform is associated with lower 12-month attrition for Tier I recruits). Conditional on remaining in the Navy for one year, no school reform has an effect on 23-month attrition (not shown).

[27] Calculation assumes that classes decrease at all grade levels. As of 1999, there were roughly 52,875,000 students and 3,304,000 teachers in public elementary and secondary schools in the United States (*Digest of Education Statistics, 2001).*

[28] Recruits from states with all four reforms are predicted to have boot camp attrition of 34 percent, whereas those from states with no reforms are predicted to have boot camp attrition of 13 percent.

Table 9—Individual results—How educational reforms affect recruits' performance in the Navy

	Bootcamp attrition		12-month attrition		Rated		Rated, Technical Field	
Sample→	Tier I	Tier II-III	Tier I	Tier II-III	Tier I	Tier II-III	Tier I	Tier II-III
Charter Legislation	0.02	0.53*	0.28**	-0.14	0.09	0.17	0.15*	0.14
	(.11)	(.31)	(.10)	(.33)	(.12)	(.27)	(.09)	(.22)
Court-Ordered	0.07	1.03**	0.19	0.02	0.15	0.33	0.29**	0.77**
Financial Reform	(.18)	(.42)	(.15)	(.37)	(.17)	(.48)	(.14)	(.34)
Test to Graduate	0.05	-0.06	0.14*	0.50**	-.18**	-0.17	-0.21**	-0.33**
	(.06)	(017)	(.07)	(.17)	(.10)	(.21)	(.07)	(.16)
Systemic Reform	-0.09	-0.16	-0.21**	0.33**	-0.24**	-0.26	-0.07	-0.11
	(.06)	(.12)	(.08)	(.13)	(.09)	(.21)	(.07)	(.18)
F-statistic[a]	7.91	9.49	22.86	30.78	8.78	2.59	11.15	8.32
	[0.10]	[0.05]	[0.00]	[0.00]	[0.07]	[.63]	[.02]	[.08]
State Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effect	Y	Y	Y	Y	Y	Y	Y	Y

Unless noted otherwise, standard errors are in parentheses.
All regressions in table 9 are estimated using logit models; because of the nonlinear nature of these models, the reported coefficients are not equal to the marginal effects. Instead, marginal effects are calculated using predicted probabilities for the entire sample.
* 10% level of significance ** 5% level of significance.
[a] Chi-squared statistic [p-value] on the joint significance of the reform variables.

Rating

When we look at the probability that a Sailor qualifies for a rating within 23 months of entering the Navy, we find that educational reforms have little effect.[29] When we examine the probability of qualifying for a technical rating, we find somewhat different results (both sets of results are shown in table 9). First, there is a general pattern across the regressions of Tier I and Tier II-III recruits; in both cases, charter legislation and court-ordered reform are predicted to increase the likelihood of qualifying, whereas test to graduate is predicted to decrease the probability of qualifying.

[29] Coming from a state with systemic or court-ordered financial reform is associated with a lower probability of qualifying for a rating for Tier I recruits, but these effects are small.

Promotion

In general, the effects of education reform on promotion (not shown) are small and insignificant. The exception to this rule concerns promotion of Tier II-III recruits to E-3 in 12 months. In this case, all four reforms together are predicted to dramatically *decrease* the probability of achieving promotion to E-3 within 12 months. (Recall that one of the most dramatic changes between 1989 and 1999 was that, by 1999, a much higher proportion of Sailors were promoted to E-3 within 12 months of joining the Navy). However, no single coefficient achieves statistical significance. Therefore, we conclude that the effects of school reform on promotion in the Navy are small.

Conclusions

Overall, the aggregated results suggest that school reforms had no influence on all test-takers and very little influence on the state-level average AFQT scores of those who joined the Navy. Of the four education reforms we examine, only test to graduate resulted in a marginal small decrease in the quality of Navy recruits.

When we examine individual-level data, we find somewhat different results: reforms seem to have a small, positive impact on AFQT scores of Sailors who completed high school and a small, negative effect on Sailors who did not graduate. To some extent, this pattern holds for the math knowledge portion of the AFQT. In terms of other outcomes, for Tier II-III recruits in particular, some reforms are associated with increased attrition. Those from states with tests to graduate are less likely to be rated in a technical field, but court-ordered financial reforms raise this likelihood for both Tier I and Tier II-III recruits.

Our data indicate that the intensive school reforms of the 1980s and 1990s were associated with very small changes in Sailors' AFQT scores. The same is generally true of on-the-job performance measures. Indeed, especially for Tier II-III recruits, AFQT scores increased fairly dramatically during this period, but the increase is not associated with any of the reforms detailed here. This increase may result from more selective recruiting on the part of the Navy and/or from a small general increase in achievement. Thus, judging from these results, even dramatic future educational reforms are unlikely to substantially change the quality of Navy recruits, or the manner in which these recruits perform on the job.

Directions for additional research

Several issues touched on here may be worthy of further exploration. In particular, among the reform variables, test to graduate had the most consistent (negative) impact on recruits, as evidenced by small declines in AFQT scores in the analyses of state-level and individual data, and declines in science subtest scores in the state-level data. In addition, test to graduate was associated with increases in boot camp attrition for both Tier I and Tier II-III recruits. It is possible that these measured effects are picking up other factors. More precise indicators of test to graduate could improve these estimates.

Perhaps of greater consequence than these effects, however, is the impact that test to graduate may have on the size of the pool of eligible recruits. As states move in the direction of policies that require passing one or more state exams to qualify for a high school diploma, and furnishing "certificates of completion" to those who stay in school but fail the state exams, it will be important for the services to evaluate the attrition patterns and performance of certificate holders. This issue is of particular importance because OSD and the services have set policies regarding the mix of high school graduates to nongraduates they will accept into military service. OSD requires 90 percent of new recruits to hold a high school diploma, and individual services have set requirements for the proportion of recruits who must be high school graduates even higher. Moreover, it is estimated that over half of public school students live in states that now have tests to graduate. In 5 years, 70 percent of public school students are expected to be subject to such exams (Center for Education Policy, 2003). Thus, it is important to reconsider recruiting policies in light of changing state requirements for high school graduation. Without a change in recruiting policy, certificate holders may be excluded from military service when, in fact, they may perform more like Tier I than Tier II-III recruits.

Second, the analyses presented here are not broken out by racial or ethnic group. If education reform works as it is intended, it should bring up the performance of students (often minority) from the weakest schools and districts. From SSI to financial reform, from competition to quality control, these are reforms that are designed to target improvement or give options to students in underresourced or failing schools. Although our descriptive statistics do not suggest obvious changes in outcomes across demographic groups, disaggregating the individual data by demographic group might yield interesting results.

Finally, State Systemic Reform is generally understood to be the weakest of the systemic reform programs, so it is not surprising that SSI had little impact on the models we analyzed. The more potent urban systemic reform could be examined; however, doing so would require additional data development in order to link recruits not only to the state they are from but, more specifically, to urban areas that held USI grants.

Appendix A: Complete regression results—state-level regression explaining AFQT scores

Sub-sample→ Variable:	All ASVAB Test-takers	Navy Accessions
Male [c]	-12.07	-0.203
	(7.81)	(6.75)
Black [c]	-20.98	-11.03
	(8.36)	(5.23)
White [c]	2.00	7.80
	(4.06)	(3.00)
Hispanic [c]	-7.78	2.64
	(5.31)	(7.53)
Navy propensity [c]	0.553	~
	(4.22)	
HS degree [c]	-11.33	-0.086
	(5.66)	(5.34)
(log of)proportion taking test/accessions [c]	-1.85	-0.876
	(0.67)	(0.492)
Poverty rate [b]	0.012	0.082
	(0.059)	(0.104)
State unempl. rate [b]	0.149	-0.183
	(0.109)	(0.222)
Proportion with college ed [b]	0.011	0.114
	(0.069)	(0.11)
Proportion metro [b]	-0.186	-0.143
	(0.225)	(0.177)
Per-pupil expenditure [b]	0.0002	0.00
	(0.0003)	(0.00)
Proportion of spending local [b]	-0.0004	0.171
	(0.129)	(0.217)
Proportion of spending state [b]	-0.005	0.161
	(0.136)	(0.22)
Teachers with MA/MS degree [b]	-0.039	-0.026
	(0.024)	(0.028)
Teachers with experience [b]	-0.038	-0.053
	(0.028)	(0.035)
Pupil-teacher ratio [b]	-0.258	-0.368
	(0.038)	(0.183)
Seats per high school Seniors [b]	0.002	0.000
	(0.009)	(0.011)
Average tuition (state college) [b]	-0.00005	0.001
	(0.0003)	(0.000)
Charter legislation [a]	-0.211	0.636
	(0.203)	(0.405)

Court-ordered financial ref [a]	0.523	0.371
	(0.527)	(0.718)
Test to graduate [a]	0.296	-0.791
	(0.308)	(0.457)
Systemic reform [a]	-0.332	-0.437
	(0.218)	(0.362)
Year = 1989 [a]	0.511	-1.41
	(0.679)	(1.16)
Year = 1994 [a]	3.22	1.78
	(0.541)	(0.91)
Constant	101.23	49.08
	(21.96)	(26.04)
Number of obs	150	150
R-squared	0.9898	0.9727

Notes: Standard errors are in parentheses.
a: dummy variable b: state-level average c: state-level averages of individual level sample
Models also include state-level fixed effects.

Appendix B: Complete regression results—individual-level regression explaining AFQT scores

Sub-sample→ Variable:	Tier I	Tier II-III
Male [a]	-0.382	0.8864
	(0.3642)	(0.6912)
Black [a]	-16.01**	-4.762**
	(0.4833)	(0.4663)
Asian-Pacific Islander [a]	-6.511**	-2.530**
	(1.160)	(1.015)
American Indian[a]	-1.463	-0.844
	(1.138)	(0.8178)
'Other' [a]	-5.272**	-2.348**
	(0.901)	(1.153)
Hispanic [a]	-6.354**	-0.970**
	(0.849)	(0.437)
Aged nineteen [a]	-3.531**	-0.5924**
	(0.1736)	(0.2506)
Aged twenty [a]	-1.858**	-0.0767
	(0.2455)	(0.3410)
Aged twenty-one or more [a]	-.06539**	0.3501
	(0.2700)	(0.5535)
High school certificate [a]	-6.680**	0.1183
	(0.8342)	(0.3025)
Some college [a]	-2.092*	~
	(1.1270)	
(log of) proportion taking test [b]	-1.705**	-1.348
	(0.7816)	(1.736)
Poverty rate [b]	0.2626**	-0.0968
	(0.0903)	(0.2022)
State unempl. rate [b]	-0.4471**	0.00737
	(0.2165)	(0.4816)
Proportion with college ed [b]	0.1095	-0.2361
	(0.1304)	(0.2613)
Proportion metro [b]	-0.0888	-0.4484*
	(0.1502)	(0.2678)
Per-pupil expenditure [b]	0.001186**	-0.000054
	(0.000473)	(0.000773)
Proportion of spending local [b]	0.5096**	0.9707**
	(0.2018)	(0.3810)
Proportion of spending state [b]	0.4637**	1.0018**
	(0.2069)	(0.3825)

Teachers with MA/MS degree [b]	-0.0420*	1.0018
	(0.0223)	(0.3810)
Teachers with experience [b]	-0.0305	-0.03402
	(0.0263)	(0.0682)
Pupil-teacher ratio [b]	-.3421**	-0.3561
	(0.1164)	(0.23388)
Seats per high school Seniors [b]	-0.00125	0.00870
	(0.00957)	(0.02528)
Average tuition (state college) [b]	0.0016**	0.0000363
	(0.00037)	(0.000972)
Charter legislation [a]	1.1936**	0.01902
	(0.4564)	(1.2708)
Court-ordered financial ref [a]	1.5084**	-0.9344
	(0.6721)	(1.8067)
Test to graduate [a]	-0.6530*	-2.05889**
	(0.3507)	(0.978)
Systemic reform [a]	-0.5254	0.2083
	(0.4497)	(0.9435)
Year = 1989 [a]	0.06698	-3.654
	(0.9433)	(1.8334)
Constant	25.535	23.168
	(18.324)	(29.351)
Number of obs	82744	9120
R-squared	0.1293	0.0354

Notes: Standard errors are in parentheses.
* 10% level of significance ** 5% level of significance.
a: dummy variable b: state-level average
Models also include state-level fixed effects.
Excluded categories: white (non-Hispanic), aged eighteen or less, high school diploma. See notes that follow for further explanation of variables.

Appendix B, continued: Descriptions of variables used in individual-level regressions

- Individual demographic controls:
  - Gender: dummy variable indicating individual is male.
  - Race: dummy variables indicating individual is African-American, Asian-Pacific Islander, American Indian, Hispanic, or 'Other'; excluded category is (non-Hispanic) Caucasian.
    - All variables from Navy's Enlisted Master Record (EMR); when EMR information is missing, data from Defense Manpower Data Center (DMDC) files.
- State-level demographic controls:
  - Proportion of the state population living in poverty (Statistical Abstracts, various years).
  - State-level unemployment rate (BLS website).
  - Proportion of the state with (at least) some college education (Statistical Abstracts, various years).
  - Proportion of the state residing in a metro area (Statistical Abstracts, various years).
  - Natural log of the (number of AFQT test-takers/1000 males aged 18-24).
- Education controls: (All data from Digest of Education Statistics, various years).
  - Total per-pupil spending.
  - Proportion of total spending from local sources.
  - Proportion of total spending from state sources.
  - Proportion of teachers with (at least) a Masters' degree.
  - Proportion of teacher with 3-9 years of experience. This measure captures the proportion of teachers who are neither very inexperienced nor nearing retirement.
    - All data from Digest of Education Statistics, various years.
- Education reform:
  - Law permitting charter schools.
  - Existence of court-ordered financial reform.
  - State requires that students pass standardized test to graduate (in addition to other requirements).
  - State had systemic reform initiative.
    - For details on sources, see table 1.
- Additionally, except where noted, specifications include year and state dummy variables (specifications are fixed-effects models).

Bibliography

Augenblick, John, J. Myers, and A. Berk Anderson, "Equity and Adequacy of School Funding," in *The Future of Children Financing Schools* Vol. 7 (3) Winter 1997.

Berger, Mark, and Eugenia Toma. 1994. "Variation in State Education Policies and Effects on Student Performance." *Journal of Policy Analysis and Management* 13(3): 477-491.

Berliner, David C. and Bruce J. Biddle. 1995. *The Manufactured Crisis.* Reading, MA: Addison-Wesley Publishing Company.

Betts, Julian. R. 1995. "Does School Quality Matter? Evidence From the National Longitudinal Survey of Youth." *Review of Economics and Statistics* 77(2): 231-50.

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100(1): 1-40.

Center on Education Policy. 2003. "State High School Exit Exams, Put to the Test."

Cooke, T. W., and A. O. Quester. 1988. *Who Stays and Who Leaves? Identifying Successful Navy Recruits,* Center for Naval Analyses CRM 2788007500.

Evans, William N., Shelia E. Murray, and Robert M. Schwab. 1997. "Schoolhouses, Courthouses, and Statehouses after Serrano." *Journal of Policy Analysis and Management* 16(1): 10-31.

Garcia, Federico E., James L. Gasch, Jennie W. Wenger, and Brian Ray. 2001. *Evaluation of the Home School and National Guard Youth ChalleNGe Program Recruits*. Center for Naval Analyses CRM D0004598.A2.

Gould, Stephen J. 1981. *The Mismeasure of Man.* New York: W. W. Norton & Co.

Grissmer, David W., Ann Flanagan, Jennifer Kawata, and Stephanie Williamson. 2000 *Improving Student Achievement: What State NAEP Test Scores Tell Us.* Santa Monica, CA: RAND, MR-924-EDU.

Hanushek, Eric A. 1989. "The Impact of Differential Expenditures on School Performance," *Educational Researcher* 18(4): 45-51.

Herrnstein, Richard J. and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Jencks, Christopher, and Meredith Phillips. 1998. *The Black-White Test Score Gap* Washington, D.C.: Brookings Institution Press.

Krueger, Alan, and Diane Whitmore. 2001. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." *Economic Journal* 11: 1-28.

Kwiatkowski, Michael. 1999. "Debating Alternative Teacher Certification: A Trial by Achievement" in *Better Teachers, Better Schools,* ed. Marci Kanstoroom and Chester E. Finn, Jr. Washington, D.C.: Thomas B. Fordham Foundation.

National Committee on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform.* http://www.ed.gov/pubs/NatAtRisk/ accessed: 25 February 2002.

Neal, Derek, and William R. Johnson. 1996. "The Role of Pre-Market Factors in Black-White Wage Differences." *Journal of Political Economy* 104: 869-95.

O'Day, J. A., and M. S. Smith. 1993. "Systemic Reform and Educational Opportunity," In S. H. Fuhrman, (Ed.), *Designing coherent education policy: Improving the system.* (pp. 250-312). San Francisco: Jossey-Bass.

Quality Counts '99: School Climate Conducive to Learning. 1999. Education Week on the Web. http://www.edweek.org/sreports/qc99/states/indicators/in_t4c.htm accessed 19 September 2000.

Rodgers, William III, and William E. Spriggs. 1996. "What Does AFQT Really Measure: Race, Wages, Schooling and the AFQT Score." *The Review of Black Political Economy* 24(4):13-46.

Rouse, Cecilia E. 1998. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113(2): 553-602.

Smith, M., and J. O'Day. 1991. "Systemic School Reform," in S. Fuhrman & B. Malen, (Eds.) *The Politics of Curriculum and Testing,* (pp. 233-267) Bristol, PA: The Falmer Press.

Tyack, David, and Larry Cuban. 1995. *Tinkering Toward Utopia: A Century of Public School Reform*. Cambridge: Harvard University Press.

US Charter Schools, 2000. "Charter School Information" (U.S. Department of Education). http://www.uscharterschools.org accessed 20 October 2000.

Witte, John F., Christopher A. Thorn, Kim M. Pritchard, and Michelle Clairborn. 1994. *Fourth-Year Report: Milwaukee Parental Choice Program.* Madison, WI: The Robert La Follette Institute of Public Affairs, University of Wisconsin-Madison.

List of tables