

**The devil's in the details:
Evidence from the GED on the large effects of
small differences in high stakes exams**

John H. Tyler^{a, *}
Richard J. Murnane^b
John B. Willett^c

a. Education Department, Box 1938, Brown University, Providence, RI 02912, USA

b. Graduate School of Education, Harvard University, Gutman Library, Cambridge, MA 02138
USA

c. Graduate School of Education, Harvard University, Gutman Library, Cambridge, MA 02138
USA

* Corresponding author. Tel: +1-401-863-1036; fax: +1-401-863-1276; E-mail address:
john_tyler@brown.edu (J. H. Tyler)

Abstract

As part of standards-based educational reform efforts, more than 40 states will soon require students to achieve passing scores on standardized exams in order to obtain a high school diploma. Currently, many states are struggling with the design of their examination systems, debating such questions as which subjects should be tested, what should be the minimum passing scores, and what should be the rules regarding options for failing students to re-take the exams. In this paper, we use data from a long-standing examination system, the General Educational Development (GED) exams, to illustrate the importance of the answers to these questions to both the number and composition of passers. We show that, among school dropouts in Florida and Texas who attempted the GED examinations, the option of retaking the examination upon initial failure was especially important for dropouts of color, and that the mathematics exam posed the greatest obstacle to females and the writing exam to males. We also show that an increase in the passing standard had a marked short run impact on the number of school dropouts in Texas who attempted the GED exams, but that the number of test-takers recovered over several years.

J.E.L. Classification I20, I28

Keywords: Efficiency, Human Capital

**The devil's in the details:
Evidence from the GED on the large effects of small changes
in examination system details**

1. Introduction

As part of standards-based educational reform efforts, more than 40 states will soon require students to achieve passing scores on standardized exams in order to obtain a high school diploma. Currently, many states struggle with the design of their examination systems, debating such questions as the number of subjects to be tested, the minimum passing scores, and the rules regarding opportunities for students to re-take the examinations. To date there is little systematic information on the impact of alternative options on who will meet the standards for a high school diploma. Nor is much known about whether particular design options will have greater impacts on some students, such as students of color, than on other students.

In this paper we use data from a long-standing examination system, the General Educational Development (GED), to examine these questions. For two reasons, the GED examination system is well suited to studying the effects of different options in state “exit tests” on the pool of students who obtain a high school diploma. First, like state exit exams, the GED exams are high stakes tests—in both cases, the results matter to the futures of the individuals taking the tests. Second, the GED exams test students who are very much like the students with whom we are often most concerned when considering the effects of state exit exams—relatively low skilled students who are on the margin of dropping out of school.

To study this population, we use GED examination data on young GED candidates from two large states to demonstrate that seemingly small changes in examination system details can have large effects on the size and composition of the passing pool. We show that the ability to retake a high stakes examination is especially important in increasing the percentage of blacks and Hispanics who pass. We also demonstrate that the “weights” attached to different subject areas in an examination system would likely alter the gender composition of the passing pool, and that seemingly small changes in the passing algorithm can drastically alter the percentage who meet the passing standard. Finally, we show that raising the passing standard results in a quite

dramatic short-run decline in the size of the test-taking pool, but that the size of the pool recovers over the next several years.

Our results are obtained using data on dropouts in Florida and Texas who were 16-19 years of age when they first attempted the GED exams. We chose this age span to make the GED testing group as similar as possible to groups of students struggling to pass state-imposed high school exit exams. The data from Florida cover GED candidates from 1991-1998, while the data from Texas cover the years 1994-2001. As we explain below, each data set has particular strengths that we exploit in answering our research questions.

2. The GED credential and examination system

During the late 1930s and early 1940s millions of young Americans left high school to serve in the military. When the veterans returned to this country, they needed a high school completion credential to be eligible for the post-secondary education benefits in the G.I. Bill. Few wanted to return to their high school classrooms. So, with governmental support, the American Council on Education established a system of examinations designed to assess whether returning veterans possessed the skills and knowledge expected of high school graduates. The examinations became known as the General Educational Development exams and those veterans who achieved passing scores were awarded the GED credential.

In 1947 New York became the first state that allowed non-veterans to obtain a high school completion credential by taking the GED exams. The other states soon followed (Auchter, 1999). The 1960s mark the beginning of a period of extraordinary growth in the number of people taking the GED examinations (Cameron and Heckman 1993b). In 2001 more than 945,000 Americans who lacked a high school diploma took the GED exams, with more than 500,000 achieving passing scores (GED Testing Service, 2002). Out of the expected half-million individuals who will drop out of school this year (U.S. Department of Education 2001), about one in three will eventually earn a GED (Murnane, Willett, and Tyler 2000). The growth in the use of the GED as an alternative high school credential that began in the 1960s has been accompanied by a substantial influx of state and federal monies into adult education enterprises

that prepare individuals to take the GED exams.

The increased profile of the GED has not escaped the attention of policy analysts and researchers. A 1992 report from the Department of Adult Education at the University of Georgia documents 55 studies examining various outcomes of GED graduates (Johnson and Valentine 1992). In the last decade there have been at least 15 additional studies looking exclusively at the economic benefits associated with GED attainment.¹ All 70 of these studies referenced above, however, are concerned primarily with the outcomes associated with GED attainment. No studies that we have found closely examine how particular details of the GED system influence the size and composition of the ultimate passing pool.

To obtain a GED, individuals must take a battery of five tests that take more than seven hours to complete. The tests cover mathematics, writing, science, social sciences, and reading. In many states GED candidates have the option to take only a portion of the entire battery at any one testing.² In all states passage is based upon some combination of the minimum test score out of the five tests and the mean score over the five tests. The GED Testing Service sets a minimum passing standard, but individual state education departments are free to set higher standards. In January 2002 the GED Testing Service introduced new versions of the GED examinations and a new scoring metric. With one exception, the results reported in this paper pertain to the tests in effect prior to 2002.

In this study we capitalize on particular features of the GED examination systems in Florida and Texas to study the effects of examination system details on outcomes. For all of the years relevant to our analysis, the GED passing standard in Florida was a minimum score of 40 on each of the five tests in the battery *and* a mean score over the five tests of 45. This is the same standard faced by GED candidates in Texas as of January 1997. However, prior to that date, Texas had a lower passing standard that required candidates to have *either* a minimum score of 40 on all five tests *or* a mean test score of 45. We use the consistent GED passing standard over time in Florida to our advantage in answering some questions, along with the change in the GED passing standard in Texas that occurred on January 1, 1997 to answer other questions. In the next section we discuss our Florida and Texas data, and then we turn to our results.

3. Data

The state departments of education in Florida and Texas provided information on the test scores and demographic characteristics of all people who took the GED examinations in those states in selected years during the 1990s. Our analytic data set for each state is composed of individuals who meet specific criteria. In Florida we used individuals who were age 16-19 when they initially tested, who first took the GED exams between the years 1991 and 1995, and whose last attempts at the GED occurred within three years of their initial test date.³ Application of these criteria provided a sample of 68,338 individuals from Florida. In Texas we used individuals who were age 16-19 when they initially tested, and who first took the GED exams between 1994 and 1998. Application of these criteria provided a sample of 143,194 from that state. Table 1 shows the distribution of initial and final test years in the Florida and Texas data.

<Table 1 about here>

As a result of the way GED data are stored in Florida, our data do not contain the complete test histories of the candidates. Rather, they contain only the initial test scores and the cumulative best scores on each test as of the last test date. Thus, in the Florida data we can determine whether individuals tested once or more than once, and we can determine how long was the period between the first and last attempt. To retain comparability, we treated the Texas data in the same way.

Table 2 presents descriptive statistics on the data from Florida and Texas. In each state the sample contains more males than females. The Florida sample has a larger percentage of white GED test-takers than the Texas sample (70 versus 46) and a smaller percentage of Hispanic test-takers (16 versus 38). The percentages of black test-takers in Florida and Texas are similar (12 and 14). In both samples, 16 year-olds are, by far, the smallest age group.⁴ In Florida the largest age group is 18-year-olds; in Texas, it is 17-year-olds. About 72 percent of the GED candidates in Florida met the passing score requirements on their initial attempt, while only 64 percent met the requirement in Texas on their initial attempt.⁵ The final GED passing rate in Florida, given the three-year attempt “window” that we allow, is 87 percent. The comparable figure for Texas is 77 percent. The average mean score in Florida for those who were eventually awarded a GED is 52, and for those who took the exam, but never met the passing standard, the mean score is 41.

The comparable mean scores for the Texas sample are 50 and 40. The remainder of the paper examines factors that contribute to the ultimate success or failure of dropouts who attempt the GED.

<Table 2 about here>

4. The role of examination system details in determining who obtains a GED

4.1. Retesting

In many states, including Massachusetts, students take exit examinations in the tenth grade. The rationale for this timing is that students who do not achieve passing scores have two years before scheduled graduation to increase their skills and to retake the exit exams. Since the option to retest is also a feature of the GED examination system, we can use the GED to examine the role of retesting in determining who might ultimately pass high stakes examinations.

Our Florida data is particularly well suited to examining the role of retesting for two reasons. First, since there were no GED policy changes in Florida over the span of our data years we can combine several years of observations without worrying about the effects of such changes on testing behavior and passing rates. Second, about 97 percent of those who attempt the GED in Florida take the complete test battery on their initial attempt. This feature is important because it means that inferences about differences in initial pass rates and ultimate pass rates reflect the role that retesting plays in affecting who passes, net of the confounding factor of strategic test-taking. By contrast, only about 73 percent of GED test takers in Texas attempt the complete battery on the first attempt. Also, Texas changed its minimum passing score requirements as of January 1997, adopting the standard Florida had used throughout the 1990s. For these reasons, we focus primarily on Florida in discussing the consequences of the retesting option. However, to provide a basis of comparison, we also provide results on the consequences of the retaking option for individuals in Texas who first attempted the GED exams in 1997 or 1998.

Panel A of Table 3 provides information on initial and ultimate GED passing rates in Florida for those who initially tested between 1991 and 1995 and took all five tests. The first row indicates that the option to retest proved to be an important facet of the GED system for many

GED candidates: the ultimate passing rate (88 percent) was 14 percentage points higher than the initial passing rate (74 percent).⁶ This overall increase masks important racial/ethnic differences in the importance of retesting. As subsequent rows of Panel A indicate, black GED candidates in Florida had an initial passing rate that was 26 percentage points below that of white candidates. However, this passing rate gap closed to a 17 percentage point gap after failing individuals retested. Hispanic testers closed an almost 9 point white-Hispanic gap by almost one-half.⁷

<Table 3 about here>

Panel B of Table 3 provides comparable information for individuals in Texas who first took all five GED exams in 1997 or 1998 (years when the passing standards were the same as those in effect in Florida) and who attempted all of the tests in the battery on their initial test date.

The patterns are similar to those in Florida. Overall, the ultimate passing rate of 80 percent was 12 points higher than the initial pass rate of 68 percent. As in Florida, the difference between the initial passing rate and the final passing rate was larger for Black test-takers (46 percent versus 64 percent) and for Hispanic test-takers (62 percent versus 76 percent) than it was for white test-takers (77 percent versus 87 percent).

To explain why the retesting option was especially important for black GED test-takers, and to a lesser extent for Hispanic test-takers, consider the following comparison. Of every 100 white dropouts who took the GED exams on their initial attempt in Florida, on average, 78 passed on their first attempt. Of the 22 who did not pass, 16 (74 percent of those who failed on initial attempt) retook the exams. Of the 16, 12 (77 percent of the test re-takers) ultimately achieved a passing score. Thus, for white GED test-takers, the option to retake the exams increased the pass rate from 78 percent to 90 percent. Of every 100 black GED test-takers in Florida, 52 passed on the first attempt. Of the 48 who did not pass, 33 (68 percent of those who failed on initial attempt) re-took the exams. Of the 33, 21 (65 percent of the test re-takers) achieved passing scores. Thus, for black GED test-takers, the option to re-take the GED exams increased the pass rate from 52 percent to 73 percent. This comparison shows that differences in the initial passing rates are the primary reason why the retesting option is more important for black and Hispanic GED examinees than for white examinees. This is far more important than differences among the groups in the percentages of initial failers who re-take the exam and differences in the

percentages of re-takers who achieve passing scores.

Of course, the analysis is not definitive because we lack information on the counterfactual -- what the passing rates would have been for different racial/ethnic groups if there had been no retesting option. Nonetheless, the data at hand suggest that the option to retest is likely to markedly affect who ultimately passes state high school exit exams, with the impact especially large for minority group members. A corollary is the importance of providing students who fail the exams with a rich set of opportunities to improve their skills.

4.2. The relative importance of subject areas

In order to put in place high school exit examinations, state officials must decide which subject areas to include on the tests, and how to weight the scores on the different subjects in calculating the overall test score. Because the GED battery covers five areas that are commonly considered “core” subjects on standardized tests, the GED systems provide an opportunity to examine the role different subject areas play in the determination of who passes. It is important to point out, however, that the science, social studies, and reading tests are to a large extent all reading comprehension tests in that they all ask questions based on comprehension of short readings. For that reason it would not be appropriate to draw inferences from this paper about the difficulties marginal students would experience in passing science tests that assessed skills and knowledge in a very different way, for example, by requiring students to carry out hands-on experiments.

A study conducted by the GED Testing Service shows that the math and writing tests tend to be the tests on which GED candidates score the lowest (Baldwin 1992). While suggestive, this study does not provide conclusions regarding which of the five subject areas of the GED exams—math, reading, writing, social studies, and science—posed the highest hurdles for GED candidates who fail the exams. Neither does the earlier study provide information on whether different demographic subgroups have relatively more difficulty with different tests in the GED battery. We use the Florida and Texas data to address these questions.

The top panel of Table 4 provides information from the Florida data on which tests gave failing GED candidates the greatest problems.⁹ The first row of the table shows that among all failing GED candidates, the lowest test scores tended to be on either the math or the writing exams (about 27 and 24 percent respectively).¹⁰ Closer examination reveals substantial gender differences. Females of every racial/ethnic group who failed to earn the credential had their lowest score on the math exam. In contrast, males of every racial/ethnic group who failed the exam had their lowest score on the writing exam. Among black males in Florida who failed to the passing standard, almost as high a percentage had their lowest score on the math exam (25) as the percentage who had their lowest score on the writing exam (27).

The bottom panel of Table 4 shows that the patterns in Texas are similar, with the math exam providing the biggest hurdle for females who failed to meeting the passing standard, and the writing exam posing the greatest obstacles for males. One exception is that among black males who failed to achieve the passing standard in Texas, a higher percentage (29) had their lowest score on the math exam than on the writing exam (22).

<Table 4 about here>

The information in Table 4 suggests that to the extent that individuals who are on the margin of passing state exit exams are similar to GED candidates, math and writing are likely to pose the greatest obstacles. At-risk female students are especially likely to need to improve their math skills, and at-risk male students are especially likely to need to improve their writing skills. Many black males are likely to need to improve both math and writing skills to pass state exit exams.

4.3. The passing rules

To implement a system of high school exit exams, states must decide what combination of scores on the exams students must achieve to graduate. Aspects of the decision include the minimum passing score on each exam, whether students must achieve passing scores on all exams, and whether students must achieve a minimum average score across the exams. The Florida and Texas GED data provide an opportunity to explore how sensitive passing rates are to different aspects of passing rules and whether the sensitivity varies across the two states.

All state passing standards include a combination of a minimum score on each of the five exams in the GED battery and a minimum mean score over the battery. In all cases the scores in each battery are given equal weight in computing the average.¹¹ In Florida individuals must score at least 40 on each of the exams in the battery and have a mean score of at least 45. After January 1997 Texas had the same passing standard. How does this scoring algorithm affect who passes relative to some other scoring scheme?

The top panel of Table 5 provides information on the roles the minimum test score requirement and the minimum mean score requirement play in determining who fails the GED exam in Florida. The bottom panel provides the comparable information for GED test-takers in Texas. Most individuals fail because they can neither meet the minimum score nor the mean score requirements. However, among individuals who can satisfy one but not the other of these criteria, there are substantial differences between initial and final failing attempts. Among those in Florida who fail on their first attempt, 14 percent do so because they have at least one score below 40, even though their average score is 45 or greater. Among those who retake the exam, but fail to achieve the credential, only 5 percent fail because they have at least one score below 40 and have an average score of at least 45. Nineteen percent of GED candidates in Florida who fail to achieve the credential on their first try do so because they do not achieve an average score of 45, even though they have no score below 40. In contrast, 42 percent of those candidates who retake the exams but still fail do so because they cannot meet the minimum average score requirement even though none of their scores is less than 40. Put together, these figures suggest that many students can improve on a single very low score, probably by concentrating their efforts on studying for the particular test. However, getting a single score from below 40 to above 40 seems easier for GED candidates to accomplish than achieving an average score of 45.¹²

The data from Texas displayed in the bottom panel of Table 5 supports this conclusion. While only 2 percent of Texas GED candidates who ultimately failed to receive the credential did so because they had one score below 40 even though they had an average score of at least 45, 50 percent failed as a result of not achieving an average passing score of 45 even though they scored at least 40 on each of the five tests.

The information in Table 5 allows us to make an initial estimate of the sensitivity of GED passing rates to the details of the passing rules. Suppose that Texas had retained its pre-1997 passing standard under which anyone with either a minimum of at least 40 or a mean of at least 45 passed. Assuming no behavior changes, 26.6% of those who initially failed in Texas after January 1997 would instead have passed on their initial attempt (7.8 plus 18.8). This implies that 5,792 of the approximately 21,784 candidates who failed on their first attempt to pass the GED exams in Texas in 1997 or 1998 would have passed had the passing standard not been raised. The story would be similar had Florida adopted the passing rules that Texas had prior to 1997. Again, assuming no behavioral changes, 32.9% percent of those who initially failed in Florida would instead have passed on their initial attempt (13.7 plus 19.2). This implies that 5,990 of the approximately 18,208 candidates who failed on their initial attempt in Florida would have passed had the passing standard been as lenient as Texas's passing standard before 1997. Of course, these simulations are only suggestive since there would likely have been different test-taking decisions by some Texas and Florida dropouts had there indeed been a different passing standard. For example, more low skilled dropouts may have attempted the GED. Nonetheless, these results do point out the potential sensitivity of passing rates to the details of the passing rules.

<Table 5 about here>

4.4. Setting the passing standard

A concern states face in introducing high school exit examinations is that making it more difficult for students to earn a high school diploma may increase the number of students who drop out of school.¹³ Since the GED examination system has been available to the civilian public for almost four decades, we cannot use this system to study the response to the sudden introduction of a high stakes test. However, we can use data from Texas to examine how a change in the passing rules on a high stakes battery of exams affected the number of dropouts who chose to take the GED exams. The relevant change is that as of January 1, 1997, GED candidates in Texas had to have minimum scores of 40 on each exam *and* an average passing score of 45. Prior to that date, they only had to satisfy one of the two criteria to obtain a GED.

Data from the GED Testing Service (GEDTS) indicate that the number of 16-19 year-old dropouts who took the GED exams in Texas fell from 41,140 in 1996 to 32,785 in 1997. While this pattern suggests that the increase in the passing standard led to a marked decline in the number of 16-19 year-old GED test-takers in Texas (the age group most relevant for this study), looking only at data for the year immediately before and the year immediately after the passing standard change could be deceptive. The number of people who took the test in 1996 may be inflated and the number who took the test in 1997 may be deflated as a result of some test-takers moving up their planned testing date from 1997 to 1996 to take advantage of the lower standard. Also, the decline in the number of test takers could be a one-year phenomenon.

To obtain better evidence on the impact of the change in the Texas GED passing standard, we used GED Testing Service data to fit a linear regression model predicting the number of 16-19 year-old GED test-takers in Texas in each year from 1990 through 1995. As indicated by the solid line in Figure 1, the model predicting a constant annual increase of 2,249 test-takers explains well the actual pattern in the number of test-takers over the period 1990-1995. We then used this fitted model to predict the number of test-takers in subsequent years on the assumption that the passing standard had not changed. This is reflected in the dashed portion of the line. We then compared the predicted number of GED test-takers with the actual number, displayed by a square, for each year from 1996 (the last year before the change in the passing standard) through 2002 (the first year the new GED exams were in effect).

In 1996 the actual number of 16-19 year-old GED test-takers (41,140) was 1,474 greater than the predicted number. This suggests that some dropouts did move up their testing date from 1997 to 1996 to take advantage of the lower passing standard. In 1997 the actual number of GED test-takers was 9,130 below the number predicted to have taken the test had the passing standard not changed. This suggests that the increase in the passing standard did result in a marked decline in the number of 16-19 year-old GED test-takers in Texas. However, the number of GED test-takers in Texas rose once again in the four years after 1997.

Figure 1 also shows that the number of 16-19 year-old GED test-takers in Texas in 2002 (13,433) was less than 40 percent of the average number over the four year period 1998-2001. The reason is that in January 2002 the GED Testing Service introduced a new set of

examinations, reported to be more difficult than the previous set. This apparently deterred many potential GED recipients from taking the exams. It led others to move up their testing date to 2001. Nationally the number of GED test-takers in 2001 was 30 percent higher than the number in 2000.¹⁴ In Texas the increase was 13 percent. Of course, it remains to be seen what the trends will be in the number of people taking the GED exams, their demographic characteristics, and their passing rate under the new exam regime.

One might argue from the evidence displayed in Figure 1 that it was a mistake for the GED Testing Service to mandate that Texas and other states with relatively low passing standards raise their standards in 1997 and to introduce new, reportedly more difficult exams in 2002. The argument would be that the increase in the passing standard reduced markedly the number of dropouts who acquired the credential. There is another side to this argument, however. Increasing the passing standard may make the GED credential more valuable to recipients. The logic is that GED recipients who meet the higher standard will earn more in the labor market because employers will learn that they possess, on average, skills greater than the average skills of dropouts who met the pre-1997 lower standard or who passed the less difficult pre-2002 GED exams¹⁵ Tyler, Murnane, and Willett (2000b) present some evidence supporting this proposition by demonstrating that school dropouts with higher scores on the GED exams have higher labor market earnings than demographically similar dropouts with lower GED exam scores.

6. Lessons

The big lesson from this paper is that the details of high stakes examination systems have large impacts on the number of people who attempt the examinations, on the number of candidates who pass, and on the racial/ethnic composition of both the test-taking group and the passing group. The option to retest is one critical design element. In Florida 21 percent of Blacks who obtained a GED did so by retaking the battery exams after an initial failure. The numbers for Hispanics and non-Hispanic whites are 17 percent and 13 percent. The comparable numbers for Texas GED test-takers are 18 for blacks, 14 for Hispanics, and 10 for non-Hispanic whites.

A second critical design element is the disciplinary content of the exams. Females in Florida

and Texas who failed the GED exam tended to do so because of low scores on the math exam; for males, the writing exam posed the greatest obstacle. Finally, the details of the passing rules are critical. 26.6 percent of GED candidates who first took the GED exams in Texas in 1997 or 1998 and failed to obtain the credential would have passed had the passing rules been those Texas used prior to 1997.

While we believe our analysis documents that “the devil lies in the details,” we do want to point out two respects in which our analysis is limited. First, our estimates of the impacts of exam system details on the number of dropouts who took GED exams and the number who passed assume that the test-taking behaviors of candidates would not have changed had the rules been different. It is likely that changes in rules would evoke some behavioral responses. For example, had the rules in Florida been that GED candidates could take the examinations only once, some candidates would probably have waited until they were better prepared before taking the exams. Others may not have taken them at all because they never felt fully prepared. A consequence of such behavioral responses is that caution is needed in drawing inferences from the results presented in this paper. For example, we show in Table 3 that 14 percent of GED test-takers in Florida failed to obtain the credential on their first try, but did so by retaking the exams. It would be inappropriate to conclude that eliminating the retesting option would reduce the percentage of GED test-takers who obtain the credential by 14 points.

The second respect in which our analysis is limited is that we focus only on the details of examination system rules, and do not consider the effects of other related policy changes. This is important in considering the impacts of standards-based educational reforms. These reforms not only include high stakes exit examinations. They also include professional development investments to improve the quality of instruction students receive and focused efforts to improve the skills of students at-risk of failing the exit exams. It is likely that these policies will affect the number of students who take the exit exams and the number who pass them, just as the details of the examination system have these effects.

Acknowledgements

The authors thank Duncan Chaplin, Carl Kaestle, and Dean Lillard for helpful comments on earlier drafts of this paper. We gratefully acknowledge the research support provided by the National Center for the Study of Adult Learning and Literacy (NCSALL), and the Russell Sage, Smith Richardson, and Spencer Foundations. Work supported by NCSALL was supported under the Educational Research and Development Centers Program, Award Number R309B60002, as administered by the Office of Educational Research and Improvement/National Institute on Postsecondary Education, Libraries, and Lifelong Learning, U.S. Department of Education. The contents do not necessarily represent the positions or policies of National Institute on Postsecondary Education, Libraries, and Lifelong Learning, U.S. Department of Education, and you should not assume endorsement by the Federal Government.

¹ See (Bos 1995; Cameron 1992; Cameron and Heckman 1993a; Cameron and Heckman 1993b; Cao, Stromsdorfer, and Weeks 1996; Cave and Bos 1994; Garet, Jing, and Kutner 1995; Maloney 1992; Murnane, Willett, and Boudett 1995; Murnane, Willett, and Boudett 1999; Murnane, Willett, and Tyler 2000; Passmore 1987; Reder 1994; Sum 1996; Tyler, Murnane, and Willett 2000a).

² Conversations the authors have had with adult education practitioners reveal, however, that GED candidates in some local settings are not informed of the option to take only a part of the battery, while some states have a de facto policy that all tests are to be taken in one sitting. Florida is such a state.

³ We limited the data in this manner so that all individuals, regardless of when they initially tested, had the same amount of time to retest. This three-year window also more closely mimics what might be expected of individuals in a state exit exam system, as opposed to the unlimited time frame for retesting that actually exists in the GED system. Since the data we received from Florida included individuals who *last tested* in the years 1998, we set a latest *initial test* year of 1995. Since the last year for which we have test data in Texas is 2001, we set a latest *initial test* year of 1998.

⁴ The small percentage of 16 year-old testers reflects age restrictions on who may take the GED tests in Florida and Texas.

⁵ In calculating the initial passing rates, we included in the initial passing category individuals who took a subset of the five GED tests on one date, the remaining tests on a later date, and passed all five tests. Thus, the difference between Florida and Texas in the initial passing rates do not reflect the larger percentage of Texas GED test-takers who took only a subset of the tests on their initial testing date. This matters more in Texas than in Florida because, as explained later in the paper, a much higher percentage of GED applicants in Texas take only some of the five tests on their initial test date than is the case for GED applicants in Florida.

⁶ This ultimate passing rate of 80.2 percent is slightly higher than the figure of 86.5 percent reported in Table 2. The reason is that the sample used for the analyses reported in Table 3 is restricted to individuals who took all five tests on their initial test date. The definition of the sample used in the analyses reported in Table 2 does not include this restriction. The same difference in sample definitions explains the difference in ultimate passing rates reported in Table 2 and Table 3 for GED test-takers in Texas.

⁷ We did not find substantial gender differences in initial passing rates.

⁹ The sample is defined as individuals who failed to meet the GED passing standard on their initial testing. It includes the scores on this initial testing for individuals who later retook the GED exams and met the passing standard.

¹⁰ Our examination of the data indicate that among those with multiple tests having the same low score (the last column in Table 4), the great majority had math and writing with equally low score. Thus, the estimate that 51 percent of those who failed did so with either math or writing as the lowest score (27 percent math, 24 percent writing) is a lower bound estimate.

¹¹ We note, however, that the writing test on the GED takes 120 minutes, the social studies test takes 85 minutes, the science test 95 minutes, the reading test 65 minutes, and the mathematics test 90 minutes, so that there is an implicit weighting across the subjects in terms of time allocation, and, potentially, required effort.

¹² We do not find race/ethnicity or gender differences in how individuals fail. Note also, that the percentage who fail because they have not finished the battery drops considerably from first to final attempt, as we would expect.

¹³ For example, Professor Walt Haney of Boston College has argued that the imposition of exit exams in Texas has led to an increase in dropout rates among black and Latino students (Haney 2000).

¹⁴ As reported in ACE News press release (2002).

¹⁵ See Betts (1998) and Costrell (1994).