DISCUSSION PAPER SERIES

IZA DP No. 6758

Labor-Market Returns to the GED Using Regression Discontinuity Analysis

Christopher Jepsen
Peter Mueser
Kenneth Troske

July 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Labor-Market Returns to the GED Using Regression Discontinuity Analysis

Christopher Jepsen
University of Kentucky

Peter Mueser
University of Missouri-Columbia and IZA

Kenneth Troske
University of Kentucky and IZA

Discussion Paper No. 6758
July 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6758
July 2012

ABSTRACT

Labor-Market Returns to the GED Using Regression Discontinuity Analysis*

We evaluate the labor-market returns to General Educational Development (GED) certification using state administrative data. We develop a fuzzy regression discontinuity (FRD) method to account for the fact that GED test takers can repeatedly retake the test until they pass it. Our technique can be applied to other situations where program participation is determined by a score on a "retake-able" test. Previous regression discontinuity estimates of the returns to GED certification have not accounted for retaking behavior, so these estimates may be biased. We find that the effect of GED certification on either employment or earnings is not statistically significant. GED certification increases postsecondary participation by up to four percentage points for men and up to eight percentage points for women.

JEL Classification: J24, I21

Keywords: GED, high school dropouts

Corresponding author:

Peter Mueser
Department of Economics
118 Professional Building
University of Missouri-Columbia
Columbia, Missouri 65211
USA
E-mail: mueserp@missouri.edu

* We thank Kyung-Seong Jeon and Laura Coogan for excellent research assistance, and we thank the Spencer Foundation for financial assistance. We thank Tom Ahn, Paco Martorell, Richard Murnane, John Papay, Jeffrey Smith and seminar participants at the 2009 APPAM conference, the 2010 Midwest Economics Association conference, the 2010 SOLE/EALE conference, the 2010 Western Economic Association conference, the 2010 Midwest Econometrics Group conference, the 2011 Econometric Society Summer Meetings, the 2011 Southern Economic Association conference, the University of Missouri, Indiana University Purdue University Indianapolis, MDRC, and the Ohio State University for useful comments. All opinions and errors are the authors' sole responsibility.

I. Introduction

Labor-market opportunities for high school dropouts have declined substantially in recent years. Certification on the General Educational Development (GED) test provides potential benefits to dropouts. Dropouts with GED certification may be able to signal to employers that they have higher skills than the “average” dropout. Many postsecondary institutions require high school graduation or GED certification for admission to degree-seeking programs.

In this paper, we evaluate the labor-market returns to GED certification in state administrative data. We develop a fuzzy regression discontinuity (FRD) method to account for the fact that GED test takers can repeatedly retake the test until they pass it. This method estimates the local average treatment effect (LATE) for test takers around the cutoff for passing the GED. Our technique, based on the discontinuity generated by the score from multiple test scores, can be applied to other situations where program participation is determined by a score on a “retake-able” test, including the original application of regression discontinuity in Thistlewaite and Campbell (1960). Other examples of retake-able tests where this technique can be applied include civil service exams, the bar exam, votes for unionization, and licensure exams such as drivers’ licenses. Previous regression discontinuity estimates of the returns to a GED have not accounted for retaking behavior, so these estimates may be biased.

We find that the estimated effects of GED certification on either employment or earnings are generally small and not statistically significant. GED certification increases postsecondary participation in the months following certification by up to four percentage points for men and up to eight percentage points for women. Finally, the results from our preferred FRD model often differ from results of a sharp regression discontinuity (SRD) design, ignoring the ability of students to retake the test.

II. Relation to Previous Literature

II.A. Regression Discontinuity Literature

Recent research on regression discontinuity (RD) methods has provided clear guidelines for determining the validity of a potential regression discontinuity analysis (Imbens and Lemieux, 2008; Lee and Lemieux, 2010; Schochet et al., 2010). However, there is little guidance for researchers with treatment measures that violate the requirements this research has identified. One often-violated criterion for a valid RD design is that the density of the variable that determines treatment, called the running variable, be smooth on either side of the discontinuity. The violation of this condition suggests that the score may be manipulated in ways that bias estimates of impact. In our context, the RD analysis requires that the GED test score density be smooth on either side of the passing threshold. As we will show later, this condition does not hold for the score on the most recent attempt of the GED test but does hold for the earlier attempts if they are properly classified.

Our paper provides a valuable contribution to the RD literature by presenting a valid RD approach for situations where the treatment is based on a test score, and individuals can retake the test in order to improve their scores. The seminal paper on RD, Thistlewaite and Campbell (1960), is an evaluation of a merit scholarship program where retaking and general manipulation of the running variable are issues. Their Table 1 shows that the merit test score density is not smooth on both sides of the passing threshold, suggesting that their regression discontinuity results may not be valid.

Little previous work has addressed the issue of test retaking with respect to RD models. Pantal, Podgursky, and Mueser (2006) show that retaking the ACT to satisfy a scholarship criterion is a significant issue, and they use initial ACT scores as instruments in their RD analysis of a college scholarship program in Missouri. Similarly, Martorell and McFarlin (2011) use the initial remedial education scores in a FRD analysis of college remediation in Texas in order to eliminate bias due to test retaking, although their analysis does not address the methodological significance of this choice.

We also present estimates that exploit the multidimensional character of the GED certification criteria, which is based on both total and subtest scores. Although there has been increased interest in developing such models (Papay, Murnane, Willett, 2011; Reardon and Robinson, 2010), current applications, especially those using an FRD design, are very limited.

II.B. GED Literature

Early work on the GED analyzed survey data from the National Longitudinal Survey of Youth (NLSY) and High School and Beyond (HSB) survey. Cameron and Heckman (1993) show that male GED recipients have lower earnings than high school graduates in a cross-section of NLSY data. They estimate models that account for the selection that occurs because wage data are not available for nonworkers, and individuals with missing wage data are unlikely to be similar to individuals with wage data on all dimensions. Heckman and LaFontaine (2006) use more recent NLSY data (through 2000) as well as two other data sets, and they find no economic returns to GED certification. Cao, Stromsdorfer, and Weeks (1996) produce similar results for women using NLSY data as well as data from Washington State.

Murnane, Willett, and Boudett (1995, 1999) extend the work on the GED based on a panel analysis of the NLSY data, comparing male GED recipients to other high school dropouts. They include multiple years of data for each person and include either person-level random effects or fixed effects to account for person-specific correlation in unobservables. The authors find positive effects of the GED on hourly wage growth. Boudett, Murnane, and Willett (2000) use the same approach and find positive effects of the GED on annual earnings for women in the NLSY.

Murnane, Willett, and Tyler (2000) use High School and Beyond (HSB) data to allow the effect of the GED on earnings to vary by cognitive ability (as measured by 10^{th} grade math scores). Using OLS models on males in the 1992 follow-up study, they find that labor-market gains associated with the GED are concentrated among recipients with low cognitive skills. Using the same model, Murnane, Willett, and Tyler (2003) find similar results for women in the HSB data.

Tyler, Murnane, and Willett (2000a, 2000b) use administrative data from the Social Security Administration (SSA) to study the effects of the GED on earnings. Using grouped data cells (to satisfy SSA data privacy requirements), they estimate differences in mean outcomes for individuals near the passing threshold in each state, thereby exploiting differences across states and over time in passing thresholds. Tyler, Murnane, and Willett (2000a) compare mean earnings by GED score for individuals aged 16-21 who took the GED in 1989 or 1990 in New York and Florida. They find a consistent, positive association between GED certification and annual earnings for nonwhite males, white females, and nonwhite females. Tyler, Murnane, and Willett (2000b) look at GED test takers aged 16 to 21 who last took the GED in 1990 in 42 states. They use method-of-moments estimators based on estimated differences in mean outcomes. The authors find positive effects of GED certification on earnings for whites (males and females) but not for nonwhites. However, Rubinstein (2003) suggests that estimates of GED impacts based on this approach could be biased because differences across states in passing standards are endogenous.

Recent GED research on earnings has utilized administrative earnings records matched with records of GED test takers to compare GED recipients with dropouts who took but did not pass the GED. For example, Tyler (2004) uses administrative data for Florida men, and he finds positive long-run earnings effects. Lofstrom and Tyler (2008) use administrative data for Texas men, but they find no impact of the GED on earnings—identified through the state's 1997 increase in the passing standard—possibly due to the low GED threshold that existed prior to the 1997 change.

Both papers use several techniques including regression discontinuity analysis. OLS models – even with pre-GED earnings – are unlikely to capture all the relevant differences between test takers who receive the GED and test takers who do not. The student fixed effects models control for all time-invariant determinants of earnings, but they cannot capture any changes over time experienced by individuals that affect both GED receipt and earnings. Although RD models offer the potential of overcoming such problems due to unmeasured factors, the validity of RD rests on several assumptions that are not adequately considered in these analyses. In particular, we show below that students' ability to retake the test if they do not pass seriously compromises the validity of estimates based on RD models that use last test score as the forcing variable as these papers do. Lofstrom and Tyler (2008) conduct a robustness test where they limit the sample to students who take the test only once, but we show below that the sample of single test takers does not satisfy conditions for implementation of a valid regression discontinuity design.

Based on both a summary of the academic literature on the GED and their own analysis, Heckman, Humphries, and Mader (2010) argue that the test has few if any benefits in terms of labor-market outcomes. They infer that there are substantial non-cognitive differences between

GED recipients and traditional high school graduates, and they suggest that the availability of the GED may induce individuals to drop out and forgo valuable non-cognitive benefits of school attendance.

Our analysis provides several contributions to the GED literature. First, the inferences that can be drawn from HSB and NLSY are limited by a lack of recent data and small samples. Both data sets contain only information on men and women in their 20s and 30s, and Heckman, Humphries, and Mader (2010) is the only study with earnings outcomes since 2001. Each data set has a sample of roughly 300 GED recipients and 300 high school dropouts of each gender. In contrast, in our analyses we will use administrative data from Missouri for over 100,000 individuals who took the GED between 1995 and 2005. We will match these data with earnings data covering the period 1993-2008, providing us with earnings for several years before and after individuals took the GED. The extended follow-up period allows us to examine the persistence of the impact of GED certification on earnings.

Second, as noted above, the previous GED research using regression discontinuity analysis failed to account for the ability of students to retake the GED. Our analysis illustrates how estimates that do not explicitly account for retaking are not valid, and we provide a technique based on multiple GED test scores to produce valid regression discontinuity estimates.

Third, nearly all the previous work focuses on the returns of the GED for men, with the studies for women based solely on NLSY, HSB, and grouped SSA data. As discussed previously, the NLSY and HSB survey data comprise small samples of dropouts and GED recipients, and there are concerns about potential endogeneity in the models using SSA data.

III. GED Test and GED Data

Nationwide, nearly 700,000 people took the GED test in 2008, and 73 percent of these received GED certification (GED Testing Service 2009). The GED test is a seven-and-a-half hour test consisting of five subtests (reading, writing, social studies, science, and mathematics). The version of the GED introduced in 2002—and referred to as the 2002 GED—replaced the previous version, which had been in place since 1988; the next version of the GED is scheduled for release in 2012.

To obtain GED certification in Missouri, test takers must obtain a minimum score on each of the five subtests and must obtain a total test score of at least 2250 out of a maximum of 4000. Certification of high school equivalency is based on a composite which combines all subtests taken over the prior two years, i.e., each subtest score is "valid" for two years before it expires. Many individuals with scores below the required thresholds retake the test—often several times—within two years, and they often retake only certain subjects rather than retaking the entire exam.[1] Scoring of tests is done at a center outside the state, and it appears unlikely that test scores could be manipulated by local administrators.

Advent of the 2002 version of the GED test altered the certification criteria in several ways. First, the minimum permitted subtest score prior to 2002 was 400, and this was raised to 410 (missing subtest scores are coded as zeros). Further, scores from earlier versions could not be combined with the 2002 version, so students who had taken the exam prior to 2002 but had not passed it had to meet the criteria based on their scores on the new version of the test. For this reason, and also because it was widely believed that the new test version would impose higher

[1] Students can take the test up to six times in any two-year period. Subject to certain constraints, states set their own criteria for certification based on test performance, but differences between states are minor, especially since 2002. Note that a given version of the test includes multiple forms that are normed to the same scale, so when a student retakes the exam, the particular questions are different.

standards, we explore the sensitivity of our findings by estimating separate models for each time period (1995-2001 and 2002-2005).

Our basic sample consists of any individual who took the GED test for the first time in Missouri between 1995 and 2005.[2] For each individual taking the test within this period, we have access to data on the most recent ten test scores taken for each version of the test, whenever the tests were taken. We exclude individuals who have taken either version of the test ten or more times because we cannot identify the first test; there were 86 individuals excluded for this reason. We exclude individuals who took the GED test while incarcerated because their labor-market outcomes are likely constrained by their incarceration.[3] We exclude individuals with missing information on gender or race/ethnicity. Individuals who received their GED through the DANTE program, which provides state certification for tests taken by military personnel outside the state, are also excluded because test scores are only reported for program participants who received GED certification; individuals who took the GED test through the military but did not pass are not in the data. Finally, we exclude individuals who took the GED as part of Missouri's GED Option program. This program, similar to those offered in several other states, allows high school students at risk of dropping out to use the GED test to help achieve a high school diploma rather than GED certification. Descriptive statistics for the regression sample are in Appendix Table A1.[4]

---

[2] Don Eisinger, Tom Robbins, and Bill Poteet of Missouri's Department of Elementary and Secondary Education provided invaluable assistance in procuring and interpreting the GED data.

[3] Tyler (2004) also points out that GED recipients with criminal records may have different labor-market returns to a GED due to their criminal history.

[4] As discussed later, the only difference between the regression sample used in our main analysis below and the full sample used in Tables 1 and 2 and Figures 1 through 3 is that the regression sample is limited to individuals with test scores between 1500 and 3000.

Quarterly earnings in all UI-covered jobs are available as reported by employers in Missouri and Kansas to the states' unemployment insurance programs. We use data through the second quarter of 2009.

Table 1 provides a tabulation of the GED scores, and an indicator of whether the test was later retaken, for individuals taking the exam for the first time in the period of our study 1995-2005. The first observation is that the overwhelming majority of individuals in our study—nearly 80 percent—obtain a score above the total passing threshold of 2250. It is therefore important to keep in mind that a regression discontinuity design will provide a LATE impact estimate for those near the threshold, individuals whose test performance is substantially below the median.

The table also shows the proportion of the test takers who retake the test within the period of our study. The bottom line in the table (right column) indicates that only about 16 percent of the test takers take the test more than once. Previous studies using RD methods have pointed to such small proportions to justify analyses that ignore test retaking. However, the overall likelihood of retaking the test is misleading in the case at hand, since the large majority of scores that satisfy the GED passing criteria with the first test are not relevant for the RD analysis, since they are far from the passing threshold. The right hand column shows that for those who do not pass, test retaking is very common. Among those with scores in the range 2200-2240, just below the passing threshold, almost 70 percent retake the GED test, and, for those with lower scores, more than half of the initial test takers retake the test. Of those who just barely meet the threshold (those with total scores 2250-2290), more than a fifth retake the test, reflecting their need to satisfy the minimum required score on each of the five subtests.

In the analysis that follows, we will define GED certification in two ways. First, when we present basic statistics on GED certification, we measure GED certification as having received GED certification during the entire sample period, i.e. by the end of 2008. This definition is the most inclusive and avoids the challenges of reporting multiple measures of GED certification. In practice, the vast majority of people who ultimately receive certification receive it within two years of first taking the test. Second, when we look at the effect of GED certification on quarterly earnings, employment, and postsecondary education, we measure GED certification at the start of the quarter in which the outcome is measured. For example, when the dependent variable is quarterly earnings 12 quarters after the initial GED test, GED certification is measured as of the start of the 12th quarter.

Test Score: Examining Discontinuities

The discussion above makes clear that individuals whose scores are close to the passing threshold are very likely to retake the GED test, yet it is the "final" test score—obtained by combining the highest subtests taken over a two-year period—that determines GED certification. Consequently, the final test score is an obvious candidate for a conventional regression discontinuity analysis. Such an approach ignores both the fact that some individuals retake the test and that some whose scores meet the overall test score threshold do not satisfy the minimum on each of the subtest scores.

Figure 1 presents the distribution of the final test scores for individuals who took the GED test in 1995-2005. The sample of test takers is slightly different from that considered above because individuals may have taken their first test prior to this period. The vertical axis identifies the number of individuals who obtain a given test score as a proportion of the total number, so the "bin size" for density calculations is a single score (possible test scores are multiples of 10). The trend line fits a local linear regression that is based on a triangular kernel with a bandwidth covering eight scores (80 points), allowing for a potential discontinuity at the threshold 2250.[5]

Simply eyeballing the curve, we can see that the discontinuity in the density is extraordinary. The log discontinuity is close to 0.92, implying that the density to the right of 2250 is approximately two and one-half times that immediately to the left, a difference that is easily statistically significant. Even though only 16 percent of individuals retake the test, the very high retake probability for those close to the cutoff point causes a dramatic redistribution in the final score.

Given that the final test score displays a marked discontinuity, it would appear highly likely that there would be discontinuities in the values for relevant characteristics. Those who choose to retake the test would be expected to differ in various ways, causing those with scores just above the threshold to differ systematically from those below. In order to test for a discontinuity in a demographic variable *X* (which we define below), we fit a fourth order polynomial in the test score, allowing for the function to change discontinuously at 2250:

$$X = \alpha_{xl} + \alpha_{xlr} D_r + \sum_{j=1}^{4} \{\beta_{xlj} [D_l (score - 2250)]^j + \beta_{xrj} [D_r (score - 2250)]^j\} + \upsilon .$$

D_r (D_l) is a dummy variable indicating whether that score equals or exceeds (is below) the passing threshold, and *score* is the total score on the GED test. β_{xlj} and β_{xrj} are estimated coefficients that capture the relationship between the GED score and the outcome variable, and the coefficient α_{xlr} identifies the extent of any discontinuity.

5 These methods correspond to those recommended by McCrary (2008).

Table 2, column 1, provides estimates for this parameter, where the variable X is one of the following: gender (male), race (nonwhite), age, whether the test taker took the test more than once, and earnings in the quarter prior to taking the test.[6] There are several statistically significant differences. Males are slightly underrepresented above the threshold, although the difference is not quite statistically significant. The proportion of nonwhites is approximately three percentage points higher above the threshold than below, a difference that is easily statistically significant. Those just above the threshold are also slightly younger and have lower prior earnings than those just below. Finally, we see that those above the threshold are slightly less likely to have retaken the test. This reflects the fact, indicated in Table 1, that those whose first test is below the threshold have much stronger incentives to retake the test. Many of them will not achieve a score that exceeds the threshold. A countervailing force—which reduces the size of the discontinuity on this measure—reflects the fact that many individuals exceed the threshold by virtue of taking the test repeated times. It is clear that the central assumptions of the RD model are violated if we take the final test score as the continuous running variable (see Imbens and Lemieux, 2008; McCrary, 2008).

One strategy to avoid this problem would be to limit consideration to the cases in which individuals have not taken the test a second time. As noted above, Lofstrom and Tyler (2008) limit their sample in this way as a robustness check for their RD estimation approach, apparently under the assumption that this group would not suffer from the same bias. Figure 2 presents the distribution of scores for individuals who took the test for the first time in the period 1995-2005 and did not take the test a second time through 2008. The most notable observation is that a

[6] The sample in the first column is the set of individuals taking the test for the *last* time between 1995 and 2005. The second column is the subset of individuals in the first column who only take the test once. The sample in the third is the set of individuals taking the test for the *first* time between 1995 and 2005.

marked discontinuity is present just as in the final test score—in fact, measured in log form, the discontinuity size is slightly larger for this sample. This similarity indicates that the discontinuity identified in the final score is not primarily a result of the fact that the final score combines scores from previous tests. Rather, it occurs because of selection, with scores just below the threshold depleted because individuals with these scores are very likely to retake the test.

A simple alternative is to use the first test score as the continuous variable underlying GED certification. Although GED certification is not predicted perfectly by the first score, there is a strong discontinuity in the relationship between first test score and ultimate GED certification, allowing us to apply a Fuzzy Regression Discontinuity (FRD) design. The FRD design requires that the first test score display continuous relationships with all pre-existing factors that may predict GED certification and employment outcomes. Table 2 (column 3) shows that there is no discontinuity in the characteristics of individuals around this measure. Figure 3 presents the distribution of the first test score, using the same method to identify discontinuities as for the densities above. Here we see that, in contrast to the final score and the score for those taking the test only once, there is essentially no discontinuity in the density at the 2250 threshold. This measure is therefore suitable for a FRD design.

Although the analysis based on the first test provides an unbiased estimate of the impact of the GED, it ignores information implicit in later test scores. For example, of those who take the test a second time, a portion will get scores that are just above the threshold of 2250, and a portion will get scores that are just below. As with the first test, the threshold will not perfectly identify whether the test taker obtains a GED, but passing a threshold is associated with the probability of ultimately obtaining a GED degree. Hence, the full distribution of scores from second tests should satisfy the basic assumptions necessary for the FRD. The same logic allows use of later test scores as well, as long as these are properly classified by parity, i.e., the number of prior tests taken.

For those retaking the test, the threshold is only relevant for an individual who obtains a greater score than that on the prior test, and so in analyses of later tests we focus on the compound score to that point and include only an individual whose compound score after taking a given test is greater than the previous scores. We also omit all those whose prior scores are above the threshold. Although this is a selected sample of test scores, the observed distribution of scores is expected to satisfy necessary continuity conditions at the threshold. This approach could be used to provide separate estimates of program impact based on each test, but, in the present analysis, we have chosen to constrain impacts to be the same in order to improve statistical power.

IV. Applying Fuzzy Regression Discontinuity Methods

The analysis will focus on first tests taken between 1995 and 2005, and second and third tests taken within two years of taking the first test. The diagnostics suggest that the each GED test score, properly identified by parity, is not subject to manipulation or selection effects. Since those at or above the test threshold are appreciably more likely to receive GED certification than those below, these data are appropriate for a fuzzy regression discontinuity (FRD) design. In order to fix ideas, we begin by describing our formal model as applied to an analysis based on the first test. We then expand the presentation to include later tests.

The equation predicting GED certification is written:[7]

$$(1) \qquad GED = \alpha_{wl} + \alpha_{wlr} D_r + \sum_{j=1}^{p} \beta_{wlj} [D_l (T - 2250)]^j + \sum_{j=1}^{p} \beta_{wrj} [D_r (T - 2250)]^j + X\eta_w + \varepsilon \,,$$

where T is the total score on the first GED test, D_l (D_r) is a dummy indicating whether that score is below (equals or exceeds) the passing threshold, p indicates the order of the polynomial, and X is a set of covariates (earnings in four quarters prior to first GED attempt, race, year of first GED test, and quarter of the year – winter, spring, summer or fall). Units are individuals who take a first GED test, and individual subscripts are suppressed. β_{wlj} and β_{wrj} are estimated coefficients identifying the relationship of the first GED test score with GED certification, below and above the 2250 threshold, respectively. The estimated parameter α_{wlr} indicates the discontinuity at the threshold.

If we fit the same structure predicting the outcome variable, we can write:

$$(2) \qquad Y = \alpha_{yl} + \alpha_{ylr} D_r + \sum_{j=1}^{p} \beta_{ylj} [D_l (T - 2250)]^j + \sum_{j=1}^{p} \beta_{yrj} [D_r (T - 2250)]^j + X\eta_y + \upsilon \,.$$

The estimate of program impact is based on the relative size of the regression discontinuity estimated in equation (1) and that estimated in equation (2). Assuming that the discontinuity in (1) induces the discontinuity in equation (2), the impact of the program can be written:

$$(3) \qquad \tau = \alpha_{ylr} / \alpha_{wlr} \,.$$

Figure 4 provides a graph that illustrates the estimation methods underlying equations (1) and (2).[8] Here the focus is on earnings in quarter 12. The discontinuity assumed in equation (1)

---

[7] The formal model presented here follows closely from that presented in Imbens and Lemieux (2008). See also McCrary (2008) and Lee and Lemieux (2010).

[8] The figure shows the results from the specification that excludes demographics. The figure with demographics has the same pattern as Figure 4.

is clearly present in the data, confirming that those who score just above the threshold on the overall GED score are appreciably more likely to have a GED within two years. The graph for earnings does not show a discontinuity at this point, suggesting that there is little impact on quarter 12 earnings.

As Imbens and Lemieux (2008) observe (see also Hahn, Todd and van der Kaauw, 2001), the FRD can be formulated as an instrumental variables system, where the treatment variable (GED certification in our case) is instrumented with the continuous measure and dummy variables capturing the discontinuity. Equation (1) is then the auxiliary equation. The outcome variable can be fitted with the following specification:

$$(4) \qquad Y = \alpha_l + \tau \widehat{GED} + \sum_{j=1}^{p} \beta_{lj} [D_l (T - 2250)]^j + \sum_{j=1}^{p} \beta_{rj} [D_r (T - 2250)]^j + X\eta + \upsilon ,$$

where $\widehat{GED}$ is the predicted value from equation (1). Since the polynomial is of the same order in equations (1) and (4), estimates of τ based on equations (1) through (3) are numerically identical to those based on equations (1) and (4). Values away from the discontinuity have no effect on the estimate of impact τ , except insofar as they influence the estimate of the extent of the discontinuity.

In order to improve precision, we perform the analysis using up to three tests, so units of analysis are individuals by test score. The specifications for the second test and the third test are the same as that specified above, except that the score on the prior test is controlled. The two-stage system implied by equations (1) and (4) are revised as follows:

$$(5) \qquad GED = \alpha_{wkl} + \alpha_{wklr} D_r + \sum_{j=1}^{p} \beta_{wklj} [D_{kl} (T_k - 2250)]^j + \sum_{j=1}^{p} \beta_{wkrj} [D_{kr} (T_k - 2250)]^j$$
$$+ \phi_{wk1} T_{k-1} + \phi_{wk2} T_{k-1}^2 + X\eta_w + \varepsilon_k$$

(6) $$Y=\alpha_{kl}+\tau\widehat{GED}+\sum_{j=1}^{p}\beta_{klj}[D_{kl}(T_k-2250)]^j+\sum_{j=1}^{p}\beta_{krj}[D_{kr}(T_k-2250)]^j$$
$$+\phi_{k1}T_{k-1}+\phi_{k2}T_{k-1}^2+X\eta+\upsilon_k$$

where k identifies the parity of the test (e.g., k=1 is the first test, k=2 is the second test), and T_k is the score associated with test k. We have added test parity subscripts to α and β, so that they are estimated separately for each of the three tests. As noted above, the sample for T_k for k=2, 3 includes only cases where $T_k > T_{k-1} < 2250$. The specification controls for prior test score for the second and third tests, but in the case of the first test, no prior test is available, i.e., we take $\phi_{w11}=\phi_{w12}=\phi_{11}=\phi_{12}=0$. Only τ is constrained to be same across all test scores, that is, we assume that the effect of getting a GED after taking any of the tests is the same, but we recognize that crossing a threshold has a different effect on the likelihood of getting a GED depending on the parity of the test, and that the relationship between test performance and earnings above and below the threshold depends on the test parity.

The above specification allows each individual to be included in the analysis up to three times, so the dependent variable is the same when the individual case is repeated, although in each case the prediction is based on a different test score. We account for clustering on individuals in calculating standard errors. For simplicity, we report the results from the quadratic model where p=2. The results from the cubic model ($p=3$) are less precisely estimated but show a similar pattern.

As noted above, our basic sample includes individuals who first take the GED test in 1995 to 2005. We exclude test takers in 2006 through 2008 because these individuals do not have sufficient earnings and education data after their initial GED test score. In addition, the sample is limited to individuals with initial test scores between 1500 and 3000 because the observed relationship between test score and GED receipt is irregular below 1500, and because there is very little variation in GED receipt above 3000. This approach eliminated 8 percent of the cases below the threshold and 12 percent of the cases above the threshold. For the remainder of the paper, we will refer to the regression analysis sample as the full sample. In keeping with previous GED research, all regressions are estimated separately for men and women.

Three dependent variables are considered for the analysis, each measured quarterly relative to the initial GED test attempt. The first dependent variable is quarterly earnings. The second measure is employment, a dichotomous variable equal to one for individuals with positive earnings in the quarter. The final measure is an indicator of whether the individual enrolled in public postsecondary education in Missouri at any time during the quarter. Earnings and employment outcomes are available for 30 quarters after the initial GED attempt, whereas postsecondary education is available for 16 quarters after the initial GED attempt.

Table 3 presents estimates based on equation (5), the first stage of the two-stage equation, applied to quarter 12.[9] In Table 3, the dependent variable is a dichotomous variable for passing the GED test, and the model is estimated with clustered standard errors for each individual. Note that the first-stage estimates for the three second-stage outcomes (quarterly income, employment, and postsecondary education) are identical because they are all based on the same sample and the same first-stage regression. The discontinuity at the threshold for the first test is associated with a 34 percentage point increase in the likelihood that men obtain GED certification. The discontinuity increases to 47 percentage points for the second attempt and 40 percentage points

[9] The results from the first-stage equation vary from quarter to quarter because the dependent variable is GED certification as of the beginning of the quarter and because the sample size varies depending on the number of GED test attempts as of the beginning of the quarter. For brevity, the table contains the results for quarter 12 after the initial GED test. The results from other quarters show a very similar pattern.

for the third attempt. For women, the discontinuities are 30 percentage points for the first attempt, 47 percentage points for the second attempt, and 46 percentage points for the third attempt. All the discontinuity variables are significant at the one-percent level (two-sided test).

Parameter estimates for the GED impact from the basic model in equations (5) and (6), estimated separately for men and women, are in the leftmost columns of Tables 4 through 6. As mentioned previously, we consider three dependent variables: quarterly earnings, employment, and postsecondary attendance. The first two columns in each table contain the estimated impact, τ, and its standard error as identified by the discontinuity in $\widehat{GED}$ from the basic FRD. The coefficient and standard error are from a separate regression for each quarter and outcome. For example, one regression is estimated for employment in the tenth quarter after the initial GED attempt.

In Table 4, the dependent variable is quarterly earnings, where quarters are measured from 1 to 30 quarters after the initial GED test attempt. The quarter in which the individual first attempts the GED is labeled quarter 0. In the basic model, although the estimated coefficients vary from quarter to quarter, only 2 of the 30 coefficients reported are statistically significant at the five-percent level (two-sided test) for men; one is statistically significant at the ten-percent level for women. Thus, in most cases, we cannot reject the hypothesis that the GED has no effect on quarterly earnings. For men, the GED does appear to be associated with higher quarterly earnings of $230 to $240 in quarters 5 and 6 following the initial GED test. The estimates have nontrivial standard errors, especially in later quarters where the sample size is smaller because individuals who took the GED in the later years do not have earnings data from all 30 quarters.

One obvious factor that may reduce employment for GED recipients would be enrollment in postsecondary education. In unreported results, we reproduced our earnings analysis limiting the sample to those not enrolled in public postsecondary education in Missouri during that quarter. The results from this sample were nearly identical to the results reported in the tables. We also estimated effects taking the dependent variable as log earnings rather than earnings, limiting consideration to those with positive earnings in the quarter. Impact estimates in this specification were qualitatively similar to those in our base analyses.[10]

In Table 5 the dependent variable is a dichotomous variable for employment, measured as having positive earnings in the quarter. None of the basic FRD coefficients in either table are statistically significant at the five percent level (two-sided test). As with earnings, the results for employment are not sensitive to the inclusion of individuals attending postsecondary education during the quarter.

Table 6 presents results for postsecondary enrollment in public institutions in Missouri. For men, the basic model indicates that GED certification is associated with increased postsecondary enrollment of two to four percentage points in the first five quarters after the test. We also see a positive effect for men of as much as three percentage points in quarters 9 to 11. In other quarters, the effect is small and not statistically significant. For women, the effect is much larger. GED certification is associated with an increased likelihood of postsecondary attendance for the first eight quarters after the initial GED attempt. The effect size is eight percentage points in the first quarter after the test, and it declines to three percentage points in the

[10] Focusing on only those employed may induce bias in estimates, but in these analyses such biases are likely to be small since GED does not predict employment (see employment results below).

eighth quarter after the test. In subsequent quarters, the effect continues to decline, and it is not statistically different from zero.

All reported estimates suffer from sizable standard errors as is typical in IV models. As noted above, in an effort to improve the estimation equations, these results are from models that control for demographic characteristics, employment prior to taking the GED, and other factors. The exclusion of the these measures increased the standard errors by as much as one third in the quarters immediately following the first GED test, but the pattern of results were nearly identical to the reported results. Because the test changed in 2002 (and prior test scores were no longer accepted at that point), we fitted models allowing the slope of the test score on GED certification and the dependent variable to differ by period. We also fitted the full model separately for the period prior to and after the implementation of the new test in 2002. In none of these models were results substantively different from those we report.

Lee and Lemieux (2010) suggest using multiple methods, both parametric and nonparametric, for conducting regression discontinuity analysis. In response, we also fitted estimates based on a local linear regression approach using software developed by Fuji, Imbens, and Kalyanaraman (2009), which, in essence, specifies a linear regression on each side of the threshold. In this approach, the choice of bandwidth is critical. Power improves as bandwidth increases, but, if there is any nonlinearity in the relationship between the running variable and the outcome, larger bandwidths induce greater bias. Because standard formulas for optimal bandwidth were unstable, we obtained estimates for a large number of bandwidths, varying from as little as 30 points (four data points), to as much as 750 points (up to 76 data points). In order to get a sense of whether nonlinearities were biasing our results, we examined graphs of the data and the estimated functions, as well as examining how estimates varied with bandwidth. We found that the appropriate bandwidth varied across cases considered. In no case were our final results based on these analyses seriously at variance with those presented here; nor were the precision of estimates substantially greater.

Multiple-Discontinuity Design

The approach above focuses on the overall GED test score, but it ignores the fact that individuals who have scores at or above 2250 face a discontinuity based on their *subtest* scores. Furthermore, those individuals who have subtest scores that are below the subtest threshold do *not* obtain the GED if their overall scores exceed the threshold, as do those with higher subtest scores. It is possible to identify sharper discontinuities based on both the total score and the lowest subtest score, essentially generalizing the FRD design to multiple dimensions.

If we create separate variables identifying whether GED overall and subtest scores meet these two criteria, the interaction between these measures identifies individuals who receive GED certification on the basis of their initial performance. The model does not, however, conform to a sharp RD design—even if reinterpreted in two dimensions—because those who fail to meet one of the criteria may still obtain GED certification when they retake the exam. This complication also opens up the possibility that there may be multiple discontinuities, which are not present in a sharp RD design. For example, when an individual has not exceeded the *overall* score threshold, if multiple test taking cannot occur, the *subtest* threshold is irrelevant. However, given the possibility of retaking the test, a subtest threshold may well influence GED certification even when the overall score falls short because those who meet the subtest criteria will have an easier time meeting the joint criteria on future tries.

Whereas the conventional FRD (or RD) setup focuses only on properly identifying the functional form of a single variable, here the functional form is multivariate. In addition to controlling for the additive impact of the overall and subtest scores, it may be necessary to recognize that the overall score and each subtest score (not just the criteria) may interact with each over. In the specification below, we therefore include continuous interactions between the overall test and subtest scores, distinguishing scores above and below the threshold.

Combining these considerations, the specification for the equation predicting GED certification, based on the first test score, can be written:

$$
\begin{aligned}
GED = \alpha_{wl} + \sum_{j=1}^{p} \beta_{wllj}[D_{Tl}D_{Sl}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{wllj}[D_{Tl}D_{Sl}(S-c)]^{j} \\
+ \phi_{wll}[D_{Tl}D_{Sl}(T-2250)(S-c)] \\
+ \alpha_{wrl}D_{Tr}D_{Sl} + \sum_{j=1}^{p} \beta_{wrlj}[D_{Tr}D_{Sl}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{wrlj}[D_{Tr}D_{Sl}(S-c)]^{j} \\
+ \phi_{wrl}[D_{Tr}D_{Sl}(T-2250)(S-c)] \\
+ \alpha_{wlr}D_{Tl}D_{Sr} + \sum_{j=1}^{p} \beta_{wlrj}[D_{Tl}D_{Sr}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{wlrj}[D_{Tl}D_{Sr}(S-c)]^{j} \\
+ \phi_{wlr}[D_{Tl}D_{Sr}(T-2250)(S-c)] \\
+ \alpha_{wrr}D_{Tr}D_{Sr} + \phi_{w0}d_{S0} + X\eta_{w} + \varepsilon,
\end{aligned} \tag{7}
$$

where the dummy variable D_{Tl} (D_{Tr}) identifies values below (equal to or above) the cutoff on the overall score, and D_{Sl} (D_{Sr}) identifies values below (equal to or above) the cutoff on the lowest subtest score. T continues to designate the total score, and S is the lowest subtest score, with the subtest threshold c.[11] The dummy variable d_{S0} indicates that the lowest subtest score is zero.[12] The estimated coefficients β_{whkj} and γ_{whkj} (h,k=l,r) identify the slope of the relationship of GED certification with the total score and the lowest subtest score, respectively, allowing different values depending on the scores relative to their thresholds. Discontinuities are estimated by α_{whk} (h,k=l,r). The interaction term $D_{Tr}D_{Sr}$ identifies individuals who receive a GED based on the

[11] For 1995-2001, c=400; for 2002 and after, c=410.

[12] In many instances, test takers choose to skip at least one subtest. As might be expected, the linear relationship assumed for the lowest test score does not apply for scores of zero.

initial test, and therefore α_{wrr} is expected to identify a major discontinuity. The smooth interaction terms are fitted with ϕ_{whk} ($h,k=l,r$). Note that when both the total and lowest subtest scores are above their respective thresholds, the actual scores are not relevant because GED certification is certain, so coefficients β_{wrrj}, γ_{wrrj} and ϕ_{wrr} are not fitted, effectively constraining their values to be zero. The test score and subtest score functions are of order p, and we will consider p=2 (quadratic).

In fitting the corresponding outcome function, the structure parallels this closely, except that discontinuities are omitted because they are the excluded instruments used for identifying the model. The outcome equation is therefore written as:

$$
\begin{aligned}
Y = \alpha_l + \tau\widehat{GED} &+ \sum_{j=1}^{p}\beta_{llj}[D_{Tl}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p}\gamma_{llj}[D_{Tl}D_{Sl}(S-c)]^j \\
&+ \phi_{ll}[D_{Tl}D_{Sl}(T-2250)(S-c)] \\
&+ \sum_{j=1}^{p}\beta_{rlj}[D_{Tr}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p}\gamma_{rlj}[D_{Tr}D_{Sl}(S-c)]^j \\
&+ \phi_{rl}[D_{Tr}D_{Sl}(T-2250)(S-c)] \\
&+ \sum_{j=1}^{p}\beta_{lrj}[D_{Tl}D_{Sr}(T-2250)]^j + \sum_{j=1}^{p}\gamma_{lrj}[D_{Tl}D_{Sr}(S-c)]^j \\
&+ \phi_{lr}[D_{Tl}D_{Sr}(T-2250)(S-c)] \\
&+ \sum_{j=1}^{p}\beta_{rrj}[D_{Tr}D_{Sr}(T-2250)]^j + \sum_{j=1}^{p}\gamma_{rrj}[D_{Tr}D_{Sr}(S-c)]^j \\
&+ \phi_{rr}[D_{Tr}D_{Sr}(T-2250)(S-c)] + \phi_0 d_{S0} + X\eta + \upsilon
\end{aligned}
\tag{8}
$$

Estimated coefficients are analogous to those in (7). The exceptions are $\beta_{rrj}, \gamma_{rrj}$ and ϕ_{rr} in (8), for which the analogous parameters are taken to be zero in equation (7)—reflecting the fact that all individuals with such scores receive GED certification. In (8) we must capture the relationship between the scores and the outcome when the GED criteria are satisfied.

Identification comes from the fact that the function in equation (8) is mostly smooth, reflecting our belief that a continuous function will identify the relationship between test scores and earnings, whereas the function determining GED receipt in equation (7) is not. As in the case of the single-dimension FRD model introduced above, the impact estimate is identified solely by the points of discontinuity, and the model fits the other relationships quite flexibly.

As a way of increasing power, we also fit this structure using a more parsimonious functional form. First, we omit the smooth interaction terms. For those below the overall test score threshold, we also constrain the slope of the overall test score to be the same whether or not the lowest subtest score is above the threshold, i.e., we take $\beta_{wllj} = \beta_{wlrj} = \beta_{wl0j}$. Analogously, we assume that, when the subtest threshold is not met, the slope of the subtest score is the same whether or not the overall score is above the threshold, i.e., $\gamma_{wllj} = \gamma_{wrlj} = \gamma_{w0lj}$. Incorporating these changes, we write a revised version of (7) as:

$$
\begin{aligned}
GED = {} & \alpha_{wl} + \sum_{j=1}^{p} \beta_{wl0j}[D_{Tl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{w0lj}[D_{Sl}(S-c)]^j \\
& + \alpha_{wrl} D_{Tr} D_{Sl} + \sum_{j=1}^{p} \beta_{wrlj}[D_{Tr} D_{Sl}(T-2250)]^j \\
& + \alpha_{wlr} D_{Tl} D_{Sr} + \sum_{j=1}^{p} \gamma_{wlrj}[D_{Tl} D_{Sr}(S-c)]^j \\
& + \alpha_{wrr} D_{Tr} D_{Sr} + \phi_{w0} d_{S0} + X\eta_w + \varepsilon
\end{aligned} \tag{9}
$$

The reduced version of equation (8) is specified analogously, except that here we also specify that $\beta_{rlj} = \beta_{rrj} = \beta_{r0j}$ and $\gamma_{rrj} = \gamma_{lrj} = \gamma_{0rj}$.

$$Y = \alpha_l + \tau \widehat{GED} + \sum_{j=1}^{p} \beta_{l0j}[D_{Tl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{0lj}[D_{Sl}(S-c)]^j$$

$$(10) \qquad + \sum_{j=1}^{p} \beta_{r0j}[D_{Tr}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{0rj}[D_{Sr}(S-c)]^j$$

$$+ \phi_0 d_{S0} + X\eta + \nu$$

For both multidimensional models, we fit equations that make use of the first three test scores, modifying estimation equations (7)-(10) to allow separate parameters by parity of test. As described in our discussion of the one-dimensional model, we control prior test score for the second and third test, and omit retake scores that are outside the range in which the discontinuity is relevant.

The third through sixth columns of Tables 4 through 6 present results from the two models presented above, again estimated as single system. As discussed previously, the first two columns contain results from the basic model. Generally, standard errors in the full multidimensional model are quite similar to those in the unidimensional model. The reduced model presented in columns (5) and (6) yields standard errors that are substantial smaller than those in the basic model (columns (1) and (2)). The impact estimates in both multidimensional models for earnings and employment yield substantive conclusions that are almost identical. Of estimates of effects on income over the 30 quarters, as in the unidimensional model, two are statistically significant and positive for men, whereas none are statistically significant for women. Effects for employment across men and women are statistically significant in only three cases, and these are negative.

Estimates of the impact on educational enrollment (Table 6) also display greater precision for the reduced multidimensional models. For men, estimates for quarters 1-4 are similar or slightly smaller than in the unidimensional model, and neither multidimensional model yields statistically significant impacts on school attendance in any subsequent quarter. In the case of women, estimates for all models are basically consistent for the first eight quarters, but the reduced model suggests that there is a larger impact on enrollment in quarters 9-11.

To address the volatility of quarterly labor market outcomes for this low-skilled population, we also estimate reduced multidimensional models where we consider aggregate measures of earnings and employment based on one to two years of data. Table 7a contains the results from these specifications for earnings, where the dependent variable is the sum of quarterly earnings for the listed quarters. Table 7b contains the results for employment, where the dependent variable is the number of quarters employed for the one- or two-year period. In each table, the top panel is for men and the bottom panel is for women. The earnings coefficients are small and not statistically significant except for a positive effect for women in years 6 and 7, which is statistically significant at the ten-percent level. The employment coefficients are almost always negative, although they are generally small and not statistically significant. The exception is that men have a lower employment of 0.199 quarters in years 4 and 5, although this effect is only significant at the ten-percent level.

In Table 8, we look at the effect of the GED on subgroups of the population based on three demographic characteristics measured at the time of the first GED attempt: race/ethnicity, age, and highest grade completed. In addition to coefficients and standard errors, the table also includes t-statistics (in absolute value) for differences in coefficients between whites and nonwhites, age 20 or less and over 20, and grade 10 or less and grade 11 or more. The table shows that there are few differences among subgroups that are statistically significant at the ten-percent level and none significant at the five-percent level. For quarters 1-4 and quarters 21-28, nonwhite women register smaller impacts of GED on employment than white women, but nonwhite women register greater impacts on postsecondary attendance. In general, the returns to the GED do not appear to differ systematically by race/ethnicity, age, or highest grade completed.

VI. Comparison with SRD Model

As discussed previously, our FRD model differs substantially from the RD models previously estimated for the GED. Specifically, previous work estimates a sharp regression discontinuity (SRD) based on the last test attempt of the GED. Therefore, in this section, we estimate SRD models of the GED in order to compare the results between methods. The comparison of models using the same sample of GED test takers is more informative than comparing our results directly with those of previous work, which looked at different time periods and states.

To the extent possible, we fit a model corresponding to that of Tyler (2004). We limit the sample to individuals who passed the subtest requirement. This requirement is needed in order to make sure that the discontinuity is sharp. With this restriction, those who obtain a score of 2250 or higher on their last attempt receive GED certification, and those who score below 2250 do not receive GED certification. We also limit the sample to individuals whose final test score is between 2200 and 2300, which is analogous to the sample used in Tyler (2004).

We estimate the following SRD model in equation (11) below:

$$(11) \quad Y = \alpha + \beta PassGED + \gamma GEDLast + X\eta + \upsilon$$

Y is the outcome of interest: earnings, employment, or education. *PassGED* is a dummy variable for receiving a score of 2250 or higher on the final attempt.[13] *GEDLast* is the score on the last

[13] Because a GED test score is "valid" for up to two years, this score is the sum of the highest subgroup scores in the two-year period up to the last GED attempt.

GED attempt, and X is a set of covariates. This model is estimated using OLS. The main difference between the model in equation (9) and the model estimated by Tyler is that Tyler uses an average of the first and last GED score, whereas we use only the final test score.

Figures 5 and 6 allow us to compare the SRD and FRD results for men and women, respectively. The solid line and box marker in each graph contains the estimated effect τ for the FRD model, as captured by the reduced multidimensional model presented in Tables 4-6. The dashed line and triangle marker is for a sharp regression discontinuity (SRD). Estimates that are statistically significant at the five-percent level (two-sided test) are shaded in black; estimates that are significant at the ten-percent level (two-sided test) are shaded in gray; and estimates that are not significant at the ten-percent level (two-sided test) are not shaded.

The figures show that the pattern of results differ by model and gender. For men's earnings, the SRD model has earnings impacts above $200 per quarter for men in a majority of quarters starting with quarter 17, although the coefficients are not significantly different from zero at the 10 percent level. In contrast, the FRD model estimates seldom exceed $100, and are statistically significant in two quarters. For women, the SRD estimates of the earnings impact increase dramatically over the period of the study, from levels that are initially negative in excess of $200, to positive impacts exceeding $400, with most of them easily statistically significant. In contrast, none of the earnings impact estimates in the FRD are statistically significant, and most are less than $100.

The differences between models for employment are more dramatic, as illustrated by the middle panels of Figures 5 and 6. The SRD model estimates much larger positive employment impacts than the FRD model, especially for women. Except for most of the first eight quarters, the SRD results for women suggest that the GED is associated with increased employment probability, whereas the FRD model suggests, if anything, negative impacts.

The SRD and FRD models produce different patterns of results for education as well. For both men and women, the FRD model estimates decline over time and eventually become statistically nonsignificant and close to zero 12 quarters after the initial GED attempt. In contrast, the SRD results increase over time, and they produce statistically significant education effects in most quarters starting with quarter 8 for men and quarter 11 for women.

The potential difference in results by model is noteworthy given our concerns about the validity of the SRD model. We have shown that the last test score is not a valid running variable for conducting an RD analysis. Our analysis for Missouri suggests that conclusions drawn from RD models that do not take into account manipulation of the running variables—due here to a test that can be retaken—may be unreliable.

VII. Conclusion

In this paper we have demonstrated how one can apply a valid regression discontinuity approach—fuzzy regression discontinuity with multiple thresholds—to a situation where a treatment is based on a test score and individuals can manipulate the test score by retaking the test. We then use this technique to estimate the effect of the GED test, which is subject to manipulation because test takers can retake the test multiple times in a two-year period. Using the FRD methodology we find that, for persons near the threshold for passing the test, the effect of passing the GED is small and statistically insignificant. We do find a positive association between passing the GED and postsecondary enrollment of up to four percentage points for men and eight percentage points for women. Given that less than 12 percent of the population of GED test takers enrolls in postsecondary institutions in any given quarter, this impact is substantial. However, these effects decline over time, becoming insignificant after five quarters for men and after 11 quarters for women.

Our results are robust to implementing the FRD technique as a local linear model as or to an instrumental variables model. Our results are also robust to the exclusion of demographic variables as well as prior earnings. However, our results are sensitive to the choice of the FRD approach as opposed to the SRD approach. The fragility of our results to the choice of technique demonstrates the importance of ensuring that the underlying assumptions of the RD estimator are met in the data. Using SRD when not appropriate may lead to the wrong conclusion.

Our findings suggest that GED certification is not of use in helping high school dropouts escape their disadvantaged labor market status. Notwithstanding these results, the GED program could be valuable if those who study for the GED—whether or not they pass it—obtain valuable skills that improve their labor market opportunities. Unfortunately, this appears highly unlikely, since the typical GED test taker spends less than 40 hours studying for the test. Perhaps most troubling, a substantial portion of high school dropouts indicate that they dropped out because they believed it was easier to obtain a GED than complete high school (Heckman et al., 2008). Insofar as additional time in school would have benefited those who drop out, the GED may have reduced the labor market success of GED test takers. At the very least, the results in this paper lend further support to the growing consensus that the GED is simply not a substitute for a high school diploma.

References

Boudett, Kathryn Parker, Richard J. Murnane, and John B. Willett. 2000. 'Second-chance' strategies for women who drop out of school. *Monthly Labor Review* (December):19-31.

Cameron, Steven V., and James J. Heckman. 1993. The nonequivalence of high school equivalents. *Journal of Labor Economics* 11, no. 1 part 1: 1-47.

Cao, Jian, Ernst W. Stromsdorfer, and Gregory Weeks. 1996. The human capital effect of General Education Development certificates on low income women. *Journal of Human Resources* 31, no.1 (Winter):206-228.

Fuji, Daisuke, Guido Imbens, and Karthik Kalyanaraman. 2009. Notes for Matlab and Stata regression discontinuity software. Unpublished manuscript downloaded July 14, 2010 from http://www.economics.harvard.edu/faculty/imbens/files/rd_software_09aug4.pdf.

General Educational Development (GED) Testing Service. 2009. GED testing fact sheet. Downloaded from http://www.acenet.edu/Content/NavigationMenu/ged/pubs/GED_Testing_Program_Fact_Sheet_2009.pdf on September 15, 2009.

Heckman, James J., John Eric Humphries and Nicholas S. Mader, 2011. The GED. Chapter 9 in *Handbook of the Economics of Education*, vol. 3, edited by Eric A Hanushek, Stephen Machin and Ludger Woessmann. New York: North Holland/Elsevier.

Heckman, James J., John Eric Humphries, Paul A. LaFontaine and Pedro L. Rodriguez, 2008. Taking the Easy Way Out: How the GED Testing Program Induces Students to Drop Out. National Bureau of Economic Research Working Paper #14044, Cambridge, MA.

Heckman, James J., and Paul A. LaFontaine. 2006. Bias-corrected estimates of GED returns. *Journal of Labor Economics* 23, no. 3 (July): 661-700.

Hahn, Jinyong, Petra Todd, and Wilbert van der Kaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69, no. 1 (January): 201-209.

Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142, no. 2 (February): 615-635.

Lee, David S., and Thomas Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Perspectives* 48, no 2. (June): 281-355.

Lofstrom, Magnus, and John H. Tyler. 2008. Modeling the signaling value of the GED with an application to an exogenous passing standard increase in Texas. *Research in Labor Economics* 28: 305-352.

Martorell, Paco, and Isaac McFarlin. 2011. Help or Hindrance? The effects of college remediation on academic and labor-market outcomes. *The Review of Economics and Statistics* 93, no. 2 (May): 436-454.

McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142, no. 2 (February): 698-714.

Murnane, Richard J., John B. Willett, and Kathryn Parker Boudett. 1995. Do high school dropouts benefit from obtaining a GED? *Educational Evaluation and Policy Analysis* 17, no.2 (Summer):133-147.

Murnane, Richard J., John B. Willett, and Kathryn Parker Boudett. 1999. Do male dropouts benefit from obtaining a GED, postsecondary education, and training? *Evaluation Review* 23, no.5 (October):475-503.

Murnane, Richard J., John B. Willett, and John H. Tyler. 2000. Who benefits from obtaining a GED? Evidence from High School and Beyond. *Review of Economics and Statistics* 82, no.1:23-37.

Pantal, Michel-Ange, Michael Podgursky, and Peter Mueser. 2006. Retaking the ACT: The effect of a state merit scholarship on test-taking, enrollment, and retention. Paper prepared for 23rd Annual Student Financial Aid Research Network Conference, Providence, RI, June 22-24, 2006.

Papay, John P., Richard J. Murnane, and John B. Willett. 2011. Extending the regression discontinuity approach to multiple assignment variables with an illustration using high school exit examinations. *Journal of Econometrics* 161, no. 2 (April): 203-207.

Reardon, Sean F., and Joseph P. Robinson. 2010. Regression discontinuity designs with multiple rating-score variables. Working paper, Stanford University.

Rubinstein, Yona. 2003. The use of interstate variation in GED passing standards for estimating the effect of a GED degree on labor-market outcomes of high school dropouts – a critique. Unpublished manuscript, Brown University.

Schochet, P, T. Cook, J. Deke, G. Imbens, J.R. Lockwood, J. Porter, and J. Smith. 2010. Standards for regression discontinuity designs. What Works Clearinghouse (June): http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf

Thistlewaite, Donald L., and Donald T. Campbell. 1960. Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *Journal of Educational Psychology* 51, no. 6 (December): 309-317.

Tyler, John H. 2004. Does GED improve earnings? Estimates from a sample of both successful and unsuccessful GED candidates. *Industrial and Labor Relations Review* 57, no.4 (July):579-598.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2000a. Do the cognitive skills of school dropouts matter in the labor market? *Journal of Human Resources* 35, no.4 (Autumn):748-754.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2000b. Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics* 115, no. 2 (May):431-468.

Table 1: Test Performance and Test Retaking: First Time Test Takers, 1995-2005

Score Range[1]	Number	Distribution (%)	Retake (%)
0-990	1,009	1.0%	65.7%
1000-1490	897	0.9%	51.7%
1500-1740	1,410	1.5%	37.6%
1750-1990	4,787	4.9%	42.9%
2000-2090	4,223	4.4%	52.9%
2100-2140	2,798	2.9%	57.0%
2150-2190	3,423	3.5%	62.0%
2200-2240	3,946	4.1%	68.9%
2250-2290	4,398	4.5%	20.9%
2300-2340	4,879	5.0%	14.0%
2350-2490	16,343	16.9%	6.8%
2500-2740	24,967	25.8%	1.6%
2750-3000	15,171	15.7%	0.4%
3010-4000	8,632	8.9%	0.2%
Total	96,883	100.0%	16.1%

[1] Only test scores that are multiples of 10 are awarded.

Table 2: Discontinuity in Estimates for the Distribution of Test Takers' Characteristics, 1995-2005.

		Final Test Score (1)	One-Time Test Takers (2)	First Test (3)
Male	Coefficient	-0.018	-0.035	0.010
	(Standard error)	(0.012)	(0.016)	(0.011)
	t-statistic	-1.5	-2.2	1.0
Nonwhite	Coefficient	0.032	0.035	0.006
	(Standard error)	(0.010)	(0.012)	(0.008)
	t-statistic	3.3	3.0	0.8
Age	Coefficient	-1.077	-1.329	-0.340
	(Standard error)	(0.219)	(0.278)	(0.191)
	t-statistic	-4.9	-4.8	-1.8
Retake test	Coefficient	-0.029		-0.504
	(Standard error)	(0.008)		(0.006)
	t-statistic	-3.6		-80.7
Prior earnings	Coefficient	-320.50	-321.95	-53.03
	(Standard error)	(242.07)	(313.38)	(210.26)
	t-statistic	-1.3	-1.0	-0.3
	Observations	95,160	81,296	96,883

Note: Prior earnings are measured as the total earnings in the four quarters (i.e. year) before the GED attempt.

Table 3: Regression Discontinuity Equation Parameter Estimates, First Stage for Quarter 12

	Men		Women	
	Coefficient	Standard Error	Coefficient	Standard Error
First test attempt				
Discontinuity	**0.33984**	**(0.00709)**	**0.30127**	**(0.00710)**
Linear - left	**0.00190**	**(0.00005)**	**0.00250**	**(0.00005)**
Linear - right	**0.00061**	**(0.00003)**	**0.00055**	**(0.00003)**
Quadratic - left	**0.00205**	**(0.00007)**	**0.00289**	**(0.00008)**
Quadratic - right	**-0.00061**	**(0.00004)**	**-0.00056**	**(0.00004)**
Second test attempt				
Discontinuity	**0.47387**	**(0.01098)**	**0.46770**	**(0.01041)**
Linear - left	**0.00123**	**(0.00006)**	**0.00127**	**(0.00006)**
Linear - right	0.00004	(0.00003)	**0.00013**	**(0.00004)**
Previous score	**-0.00080**	**(0.00026)**	-0.00019	(0.00030)
Previous score squared	**2.0E-07**	**(6.1E-08)**	5.1E-08	(7.1E-08)
Dummy for second attempt	**0.69914**	**(0.27528)**	-0.00596	(0.32522)
Third test attempt				
Discontinuity	**0.39571**	**(0.02228)**	**0.45537**	**(0.02080)**
Linear - left	**0.00161**	**(0.00016)**	**0.00198**	**(0.00015)**
Linear - right	0.00017	(0.00010)	0.00015	(0.00009)
Previous score	0.00012	(0.00031)	**0.00102**	**(0.00026)**
Previous score squared	-2.4E-08	(8.3E-08)	**-3.0E-07**	**(7.4E-08)**
Dummy for third attempt	-0.33545	(0.30313)	**-1.03227**	**(0.24406)**
Observations	51,637		48,961	
Number of Individuals	44,378		41,967	
Adjusted R-squared	0.58		0.60	

Notes: Bold terms represent coefficients that are statistically significant at the five-percent level (two-sided test). Quadratic terms and earnings variables are measured in thousands. Standard errors are clustered by individual. Each regression also includes earnings in the four quarters prior to first GED attempt, race, dummy variables for year of first GED test, and quarter of the year (winter, spring, summer or fall).

Table 4a: Estimated GED Impact on Earnings for Alternative FRD Designs, Men

Quarters since 1st GED test	Basic		Full Multidimensional		Reduced Multidimensional		Mean Earnings
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	
1	-36.2	(66.9)	82.0	(59.1)	27.3	(42.8)	$2,048
2	71.2	(80.1)	63.9	(77.2)	5.4	(56.0)	$2,190
3	78.8	(90.8)	74.4	(87.0)	-2.4	(63.1)	$2,255
4	9.0	(103.9)	127.0	(99.4)	-5.8	(73.0)	$2,333
5	**228.2**	**(114.6)**	**281.2**	**(106.1)**	**199.3**	**(78.7)**	$2,404
6	**238.9**	**(120.6)**	**296.5**	**(114.8)**	*124.2*	*(86.2)*	$2,470
7	210.5	(131.8)	*224.1*	*(123.7)*	103.5	(92.6)	$2,507
8	133.1	(139.2)	-8.8	(131.7)	63.6	(96.7)	$2,565
9	131.7	(144.8)	118.8	(136.2)	70.0	(100.1)	$2,607
10	125.2	(155.9)	-143.0	(145.4)	-60.8	(106.3)	$2,657
11	242.2	(156.9)	-28.4	(147.0)	27.2	(108.5)	$2,661
12	70.0	(167.7)	60.1	(155.0)	35.3	(114.2)	$2,723
13	-16.7	(174.8)	-7.4	(159.0)	-97.5	(118.0)	$2,766
14	-42.4	(179.3)	-47.6	(161.9)	-98.6	(120.1)	$2,781
15	-11.8	(182.6)	77.7	(167.1)	-32.2	(123.7)	$2,795
16	-129.6	(185.2)	59.1	(169.6)	-114.8	(126.2)	$2,835
17	-11.6	(188.0)	71.7	(174.9)	-44.1	(127.8)	$2,858
18	-147.9	(194.7)	-65.9	(179.2)	-93.4	(130.5)	$2,888
19	38.7	(193.4)	-27.3	(178.0)	-13.8	(131.2)	$2,904
20	-45.9	(201.5)	-13.9	(183.7)	-74.7	(137.1)	$2,933
21	72.7	(206.7)	-61.6	(191.4)	59.5	(140.2)	$2,971
22	178.1	(218.5)	-24.6	(199.9)	77.1	(146.6)	$3,016
23	11.8	(224.6)	98.7	(207.7)	97.4	(150.3)	$3,018
24	34.8	(232.2)	-7.7	(214.3)	-66.8	(158.2)	$3,049
25	217.4	(226.2)	233.3	(214.4)	115.5	(155.5)	$3,078
26	50.6	(231.2)	303.0	(221.6)	74.5	(161.1)	$3,104
27	-32.0	(232.8)	78.6	(226.0)	-36.5	(163.8)	$3,090
28	39.3	(233.5)	190.7	(227.4)	164.1	(165.5)	$3,135
29	-51.4	(261.5)	196.1	(251.2)	110.4	(173.9)	$3,178
30	-90.0	(252.8)	-24.8	(246.3)	-28.9	(177.2)	$3,211

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 4b: Estimated GED Impact on Earnings for Alternative FRD Designs, Women

Quarters since 1st GED test	Basic		Full Multidimensional		Reduced Multidimensional		Mean Earnings
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	
1	-56.8	(59.7)	-2.8	(54.7)	29.1	(37.7)	$1,725
2	-40.5	(71.2)	-14.9	(67.9)	-17.5	(48.2)	$1,864
3	21.4	(80.2)	45.7	(80.6)	-21.2	(58.9)	$1,921
4	68.1	(89.8)	64.0	(92.6)	9.3	(62.2)	$1,996
5	2.3	(94.9)	-19.6	(98.9)	-37.5	(69.9)	$2,058
6	41.5	(102.6)	-19.3	(107.7)	-47.0	(71.6)	$2,109
7	-49.9	(112.8)	-105.4	(119.4)	-45.0	(76.6)	$2,155
8	-18.7	(116.7)	-103.7	(126.2)	-103.0	(80.7)	$2,227
9	-72.1	(125.2)	-45.7	(133.2)	-79.3	(85.2)	$2,261
10	5.6	(131.1)	-40.3	(140.8)	-70.0	(88.3)	$2,300
11	17.0	(135.8)	-18.3	(148.1)	-5.1	(93.5)	$2,312
12	37.8	(140.0)	6.5	(150.2)	-30.4	(97.0)	$2,359
13	5.6	(144.7)	67.8	(155.4)	45.0	(99.3)	$2,382
14	-7.3	(151.4)	39.3	(162.7)	39.6	(103.2)	$2,398
15	110.0	(149.4)	121.6	(157.6)	68.8	(102.1)	$2,413
16	181.7	(151.6)	76.4	(160.1)	71.5	(103.2)	$2,434
17	116.0	(159.1)	2.5	(163.8)	39.4	(107.0)	$2,457
18	102.8	(164.0)	-8.7	(167.0)	22.6	(109.5)	$2,483
19	125.0	(162.6)	-67.5	(168.7)	36.4	(109.2)	$2,473
20	211.5	(169.4)	-9.5	(170.4)	57.3	(111.7)	$2,477
21	*294.2*	*(169.9)*	-38.2	(174.4)	148.6	(115.0)	$2,500
22	267.7	(190.4)	-80.2	(197.0)	111.1	(126.0)	$2,527
23	111.4	(178.2)	-60.4	(188.0)	143.6	(122.2)	$2,517
24	204.8	(182.1)	76.6	(192.7)	97.1	(129.6)	$2,555
25	242.3	(182.1)	130.7	(191.7)	200.9	(127.1)	$2,566
26	60.1	(190.2)	7.0	(201.6)	64.1	(131.2)	$2,598
27	88.7	(191.0)	-126.8	(207.2)	68.6	(134.3)	$2,576
28	127.6	(191.5)	-34.3	(205.3)	91.6	(135.3)	$2,604
29	135.8	(200.5)	-204.3	(224.9)	132.2	(141.2)	$2,604
30	-128.9	(205.2)	-113.2	(233.0)	15.5	(143.7)	$2,633

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 5a: Estimated GED Impact on Employment for Alternative FRD Designs, Men

Quarters since 1st GED test	Basic Coeff.	Basic Std. Err.	Full Multidimensional Coeff.	Full Multidimensional Std. Err.	Reduced Multidimensional Coeff.	Reduced Multidimensional Std. Err.	Mean Employ.
1	0.013	(0.017)	0.015	(0.015)	0.006	(0.010)	62.1%
2	-0.015	(0.018)	-0.020	(0.017)	-0.015	(0.012)	62.6%
3	0.016	(0.019)	-0.002	(0.019)	-0.012	(0.013)	62.1%
4	-0.014	(0.020)	-0.010	(0.020)	-0.017	(0.014)	61.7%
5	0.000	(0.021)	-0.003	(0.021)	-0.005	(0.015)	61.7%
6	-0.008	(0.022)	0.003	(0.022)	-0.024	(0.015)	61.1%
7	-0.010	(0.023)	-0.006	(0.022)	-0.013	(0.016)	60.7%
8	0.007	(0.024)	-0.005	(0.023)	0.001	(0.016)	60.5%
9	0.004	(0.024)	-0.020	(0.024)	-0.019	(0.017)	60.0%
10	-0.009	(0.025)	-0.015	(0.024)	-0.021	(0.017)	59.4%
11	0.000	(0.025)	-0.027	(0.024)	-0.016	(0.017)	58.7%
12	-0.022	(0.026)	-0.006	(0.025)	-0.022	(0.018)	58.3%
13	*-0.045*	*(0.026)*	-0.020	(0.025)	*-0.032*	*(0.018)*	57.5%
14	-0.036	(0.026)	-0.022	(0.025)	*-0.032*	*(0.018)*	56.9%
15	-0.002	(0.026)	-0.013	(0.025)	-0.023	(0.018)	56.4%
16	-0.011	(0.027)	-0.005	(0.026)	-0.019	(0.018)	56.1%
17	-0.018	(0.027)	-0.005	(0.026)	-0.016	(0.018)	55.6%
18	-0.013	(0.027)	-0.019	(0.026)	-0.024	(0.018)	55.5%
19	-0.007	(0.027)	-0.025	(0.026)	-0.025	(0.019)	55.2%
20	0.004	(0.027)	0.001	(0.026)	-0.025	(0.019)	54.6%
21	-0.017	(0.028)	-0.041	(0.027)	-0.031	(0.019)	54.5%
22	-0.011	(0.028)	**-0.054**	**(0.028)**	-0.032	(0.020)	54.5%
23	-0.027	(0.029)	-0.010	(0.028)	-0.029	(0.020)	53.9%
24	-0.027	(0.029)	-0.016	(0.028)	-0.022	(0.020)	53.6%
25	-0.021	(0.029)	-0.001	(0.029)	*-0.038*	*(0.020)*	53.4%
26	-0.038	(0.029)	0.016	(0.029)	-0.017	(0.021)	53.3%
27	*-0.052*	*(0.029)*	-0.005	(0.030)	*-0.040*	*(0.021)*	52.8%
28	-0.033	(0.030)	0.000	(0.030)	-0.017	(0.021)	52.7%
29	-0.026	(0.030)	0.008	(0.030)	-0.006	(0.022)	52.7%
30	*-0.057*	*(0.030)*	-0.018	(0.030)	-0.022	(0.022)	52.5%

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 5b: Estimated GED Impact on Employment for Alternative FRD Designs, Women

Quarters since 1st GED test	Basic Coeff.	Basic Std. Err.	Full Multidimensional Coeff.	Full Multidimensional Std. Err.	Reduced Multidimensional Coeff.	Reduced Multidimensional Std. Err.	Mean Employ.
1	-0.005	(0.017)	-0.001	(0.016)	-0.004	(0.011)	63.1%
2	-0.016	(0.018)	-0.015	(0.018)	-0.008	(0.012)	64.1%
3	0.002	(0.020)	-0.004	(0.020)	-0.004	(0.014)	64.2%
4	-0.010	(0.021)	0.004	(0.022)	-0.003	(0.014)	64.6%
5	0.001	(0.022)	0.023	(0.023)	-0.005	(0.015)	64.4%
6	-0.025	(0.023)	0.001	(0.024)	-0.019	(0.016)	64.3%
7	-0.012	(0.024)	0.004	(0.025)	-0.016	(0.017)	64.0%
8	-0.017	(0.024)	-0.014	(0.026)	-0.024	(0.017)	64.0%
9	*-0.045*	*(0.025)*	-0.021	(0.027)	**-0.043**	**(0.017)**	63.9%
10	-0.020	(0.025)	-0.026	(0.027)	**-0.037**	**(0.018)**	63.5%
11	-0.011	(0.026)	-0.009	(0.028)	-0.015	(0.018)	62.7%
12	0.001	(0.027)	0.019	(0.028)	-0.009	(0.018)	62.5%
13	0.000	(0.027)	0.003	(0.029)	-0.010	(0.019)	62.2%
14	0.003	(0.028)	0.021	(0.029)	0.004	(0.019)	61.9%
15	0.033	(0.028)	0.034	(0.030)	0.009	(0.019)	61.3%
16	0.046	(0.029)	0.035	(0.030)	0.019	(0.019)	60.8%
17	0.041	(0.029)	0.033	(0.030)	0.007	(0.019)	60.5%
18	0.006	(0.029)	-0.009	(0.030)	-0.021	(0.020)	60.2%
19	-0.018	(0.029)	-0.019	(0.031)	*-0.035*	*(0.020)*	59.6%
20	0.036	(0.030)	-0.001	(0.031)	-0.007	(0.020)	59.3%
21	-0.002	(0.031)	-0.029	(0.032)	-0.009	(0.021)	58.8%
22	0.004	(0.031)	-0.039	(0.032)	-0.012	(0.021)	58.4%
23	-0.011	(0.031)	-0.046	(0.033)	-0.008	(0.021)	58.0%
24	-0.011	(0.032)	-0.013	(0.033)	-0.004	(0.022)	57.9%
25	0.040	(0.032)	0.009	(0.034)	0.021	(0.022)	57.5%
26	0.020	(0.032)	0.033	(0.035)	0.003	(0.022)	57.5%
27	0.000	(0.032)	-0.005	(0.035)	-0.005	(0.023)	57.0%
28	0.026	(0.033)	0.037	(0.035)	0.007	(0.023)	56.7%
29	0.011	(0.033)	0.001	(0.036)	0.004	(0.023)	56.4%
30	0.027	(0.033)	0.021	(0.037)	0.023	(0.023)	56.2%

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 6a: Estimated GED Impact on Education for Alternative FRD Designs, Men

Quarters since 1st GED test	Basic Coeff.	Basic Std. Err.	Full Multidimensional Coeff.	Full Multidimensional Std. Err.	Reduced Multidimensional Coeff.	Reduced Multidimensional Std. Err.	Mean Education
1	**0.039**	**(0.007)**	**0.038**	**(0.007)**	**0.029**	**(0.005)**	7.7%
2	**0.032**	**(0.008)**	**0.038**	**(0.007)**	**0.027**	**(0.006)**	7.2%
3	**0.022**	**(0.008)**	**0.030**	**(0.008)**	**0.022**	**(0.006)**	6.7%
4	**0.023**	**(0.008)**	**0.030**	**(0.008)**	**0.024**	**(0.006)**	6.4%
5	**0.024**	**(0.008)**	**0.023**	**(0.009)**	**0.016**	**(0.007)**	6.0%
6	*0.016*	*(0.009)*	*0.015*	*(0.009)*	0.004	(0.007)	5.5%
7	0.002	(0.009)	-0.001	(0.009)	-0.003	(0.007)	5.2%
8	0.006	(0.009)	-0.002	(0.009)	0.004	(0.007)	5.0%
9	**0.025**	**(0.008)**	0.013	(0.009)	*0.011*	*(0.006)*	4.8%
10	**0.031**	**(0.008)**	0.005	(0.009)	0.009	(0.006)	4.5%
11	**0.023**	**(0.008)**	0.002	(0.008)	0.007	(0.006)	4.2%
12	0.013	(0.008)	0.001	(0.009)	0.002	(0.006)	4.1%
13	0.004	(0.009)	-0.004	(0.008)	-0.006	(0.006)	3.8%
14	0.011	(0.008)	0.003	(0.008)	-0.001	(0.006)	3.6%
15	0.006	(0.008)	-0.002	(0.009)	-0.002	(0.006)	3.5%
16	0.002	(0.008)	-0.001	(0.008)	0.001	(0.006)	3.3%

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 6b: Estimated GED Impact on Education for Alternative FRD Designs, Women

Quarters since 1st GED test	Basic Coeff.	Basic Std. Err.	Full Multidimensional Coeff.	Full Multidimensional Std. Err.	Reduced Multidimensional Coeff.	Reduced Multidimensional Std. Err.	Mean Education
1	**0.080**	**(0.009)**	**0.077**	**(0.009)**	**0.065**	**(0.007)**	12.0%
2	**0.077**	**(0.009)**	**0.068**	**(0.010)**	**0.064**	**(0.007)**	11.7%
3	**0.056**	**(0.010)**	**0.073**	**(0.011)**	**0.057**	**(0.008)**	10.9%
4	**0.053**	**(0.010)**	**0.055**	**(0.011)**	**0.049**	**(0.008)**	10.4%
5	**0.051**	**(0.011)**	**0.051**	**(0.012)**	**0.049**	**(0.008)**	9.7%
6	**0.038**	**(0.011)**	**0.042**	**(0.012)**	**0.042**	**(0.008)**	9.2%
7	**0.035**	**(0.012)**	**0.029**	**(0.013)**	**0.041**	**(0.009)**	8.9%
8	**0.032**	**(0.012)**	**0.029**	**(0.013)**	**0.037**	**(0.009)**	8.4%
9	*0.021*	*(0.012)*	*0.023*	*(0.014)*	**0.031**	**(0.009)**	8.0%
10	0.011	(0.013)	0.015	(0.013)	**0.023**	**(0.009)**	7.7%
11	0.018	(0.012)	0.013	(0.013)	**0.024**	**(0.009)**	7.4%
12	0.014	(0.012)	-0.001	(0.013)	0.013	(0.009)	7.0%
13	0.004	(0.012)	0.001	(0.013)	0.010	(0.009)	6.6%
14	0.002	(0.013)	-0.009	(0.013)	0.008	(0.009)	6.4%
15	-0.003	(0.013)	-0.005	(0.013)	0.010	(0.009)	6.1%
16	-0.002	(0.012)	0.004	(0.013)	0.012	(0.009)	5.9%

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 7a: Estimated GED Impact for Total Earnings
Multidimensional Reduced Model

Quarters since 1st GED test	Coeff.	Standard Error	Mean Earnings	Coeff. As Pct Of Mean	Std. Err. As Pct Of Mean
		Men			
1-4	42.9	(162.9)	8,750	0.5%	1.9%
5-12	459.4	(584.3)	20,470	2.2%	2.9%
13-20	-626.2	(855.1)	22,813	-2.7%	3.7%
21-28	496.5	(1109.1)	24,594	2.0%	4.5%
		Women			
1-4	34.5	(145.1)	7,465	0.5%	1.9%
5-12	-226.3	(483.2)	17,717	-1.3%	2.7%
13-20	386.7	(706.2)	19,514	2.0%	3.6%
21-28	*1612.9*	*(913.6)*	20,532	7.9%	4.4%

Table 7b: Estimated GED Impact for Total Quarters Employed
Multidimensional Reduced Model

Quarters since 1st GED test	Coeff.	Standard Error	Mean Employ.	Coeff. As Pct Of Mean	Std. Err. As Pct Of Mean
		Men			
1-4	-0.012	(0.032)	2.461	-0.5%	1.3%
5-12	-0.123	(0.091)	4.768	-2.6%	1.9%
13-20	*-0.199*	*(0.115)*	4.472	-4.4%	2.6%
21-28	-0.188	(0.135)	4.295	-4.4%	3.1%
		Women			
1-4	-0.007	(0.034)	2.547	-0.3%	1.3%
5-12	-0.127	(0.091)	5.101	-2.5%	1.8%
13-20	-0.031	(0.121)	4.890	-0.6%	2.5%
21-28	0.054	(0.145)	4.672	1.2%	3.1%

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 8: Estimated GED Impact for Subgroups, Reduced Multidimensional Model

Quarters since 1st GED test	Nonwhite Coeff.	Nonwhite Std. Err.	White Coeff.	White Std. Err.	Diff. T-Stat.	20 or Under Coeff.	20 or Under Std. Err.	Over 20 Coeff.	Over 20 Std. Err.	Diff. T-Stat.
Men										
Total Earnings										
1-4	253	(269)	-43.1	(201)	0.9	262	(183)	-122	(283)	1.1
5-12	917	(939)	346	(721)	0.5	653	(710)	1031	(962)	0.3
13-20	516	(1418)	-999	(1046)	0.9	787	(1072)	-1517	(1388)	1.3
21-28	*3127*	*(1849)*	-123	(1343)	1.4	1417	(1407)	-50.7	(1766)	0.6
Quarters Employed										
1-4	-0.003	(0.066)	-0.020	(0.037)	0.2	0.042	(0.043)	-0.070	(0.048)	1.7
5-12	-0.258	(0.183)	-0.051	(0.105)	1.0	-0.059	(0.122)	-0.138	(0.137)	0.4
13-20	-0.181	(0.224)	-0.190	(0.133)	0.0	0.031	(0.153)	**-0.371**	**(0.173)**	1.7
21-28	-0.100	(0.269)	-0.227	(0.155)	0.4	-0.074	(0.177)	-0.242	(0.205)	0.6
Quarters Enrolled in Postscondary Education										
1-4	**0.127**	**(0.036)**	**0.061**	**(0.018)**	1.7	**0.060**	**(0.023)**	**0.095**	**(0.021)**	1.1
5-12	0.031	(0.083)	0.043	(0.038)	0.1	-0.004	(0.052)	0.074	(0.049)	1.1
Women										
Total Earnings										
1-4	-46.1	(297)	83.2	(167)	0.4	113	(183)	-42.3	(215)	0.6
5-12	-160	(958)	-355	(570)	0.2	-123	(651)	-386	(691)	0.3
13-20	-19.3	(1415)	595	(818)	0.4	248	(996)	520	(985)	0.2
21-28	840	(1848)	**2075**	**(1057)**	0.6	589	(1295)	*2177*	*(1270)*	0.9
Quarters Employed										
1-4	*-0.118*	*(0.071)*	0.026	(0.039)	1.8	-0.043	(0.051)	0.016	(0.046)	0.9
5-12	-0.251	(0.174)	-0.080	(0.109)	0.8	-0.175	(0.136)	-0.079	(0.123)	0.5
13-20	-0.063	(0.223)	-0.007	(0.145)	0.2	-0.081	(0.187)	0.017	(0.160)	0.4
21-28	-0.358	(0.279)	0.213	(0.173)	1.7	-0.100	(0.225)	0.111	(0.193)	0.7
Quarters Enrolled in Postscondary Education										
1-4	**0.273**	**(0.056)**	**0.155**	**(0.024)**	1.9	**0.131**	**(0.036)**	**0.204**	**(0.028)**	1.6
5-12	**0.424**	**(0.111)**	**0.188**	**(0.054)**	1.9	**0.195**	**(0.080)**	**0.254**	**(0.062)**	0.6

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual. Control variables are listed in the notes to Table 3.

Table 8 (Continued): Estimated GED Impact for Subgroups, Reduced Multidimensional Model

Quarters since 1st GED test	Grade 10 or less		Grade 11 or more		
	Coeff.	Std. Err.	Coeff.	Std. Err.	T-Stat.
Men					
Total Earnings					
1-4	25.4	(211.1)	534.4	(351.6)	-1.2
5-12	341.5	(773.8)	1966.9	(1258.3)	-1.1
13-20	-1059.7	(1118.4)	1079.1	(1931.3)	-1.0
21-28	-108.3	(1354.5)	3730.8	(2368.1)	-1.4
Quarters Employed					
1-4	-0.027	(0.044)	*0.109*	*(0.064)*	-1.8
5-12	-0.145	(0.123)	-0.029	(0.186)	-0.5
13-20	-0.230	(0.152)	-0.165	(0.242)	-0.2
21-28	-0.184	(0.168)	-0.154	(0.268)	-0.1
Quarters Enrolled in Postscondary Education					
1-4	**0.089**	**(0.020)**	0.026	(0.035)	1.5
5-12	0.040	(0.046)	0.053	(0.081)	-0.1
Women					
Total Earnings					
1-4	145.8	(193.8)	62.6	(304.9)	0.2
5-12	130.8	(637.9)	-456.1	(1008.0)	0.5
13-20	*1521.4*	*(912.7)*	-346.2	(1471.3)	1.1
21-28	**2829.1**	**(1104.5)**	-463.9	(1778.9)	1.6
Quarters Employed					
1-4	-0.027	(0.046)	-0.020	(0.069)	-0.1
5-12	-0.088	(0.122)	-0.286	(0.181)	0.9
13-20	0.128	(0.159)	-0.380	(0.231)	1.8
21-28	0.175	(0.178)	-0.270	(0.270)	1.4
Quarters Enrolled in Postscondary Education					
1-4	**0.161**	**(0.028)**	**0.138**	**(0.047)**	0.4
5-12	**0.234**	**(0.064)**	**0.219**	**(0.104)**	0.1

Notes: Coefficients in bold are statistically significant at the five-percent level (two-sided test), and coefficients in italics are statistically significant at the ten-percent level (two-sided test). Each coefficient is from a separate regression. Standard errors are clustered by individual.

Figure 1: Distribution of Final Test Score, 1995-2005

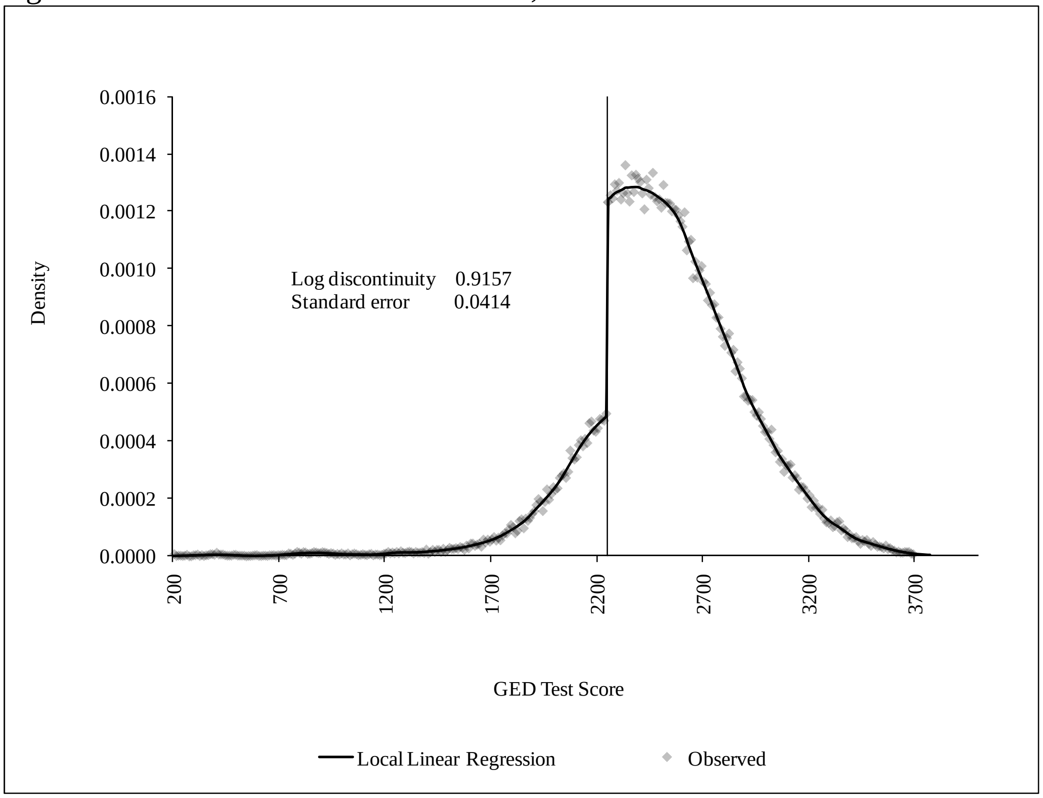


Figure 2: Distribution of First Test Score for Single Test Takers, 1995-2005

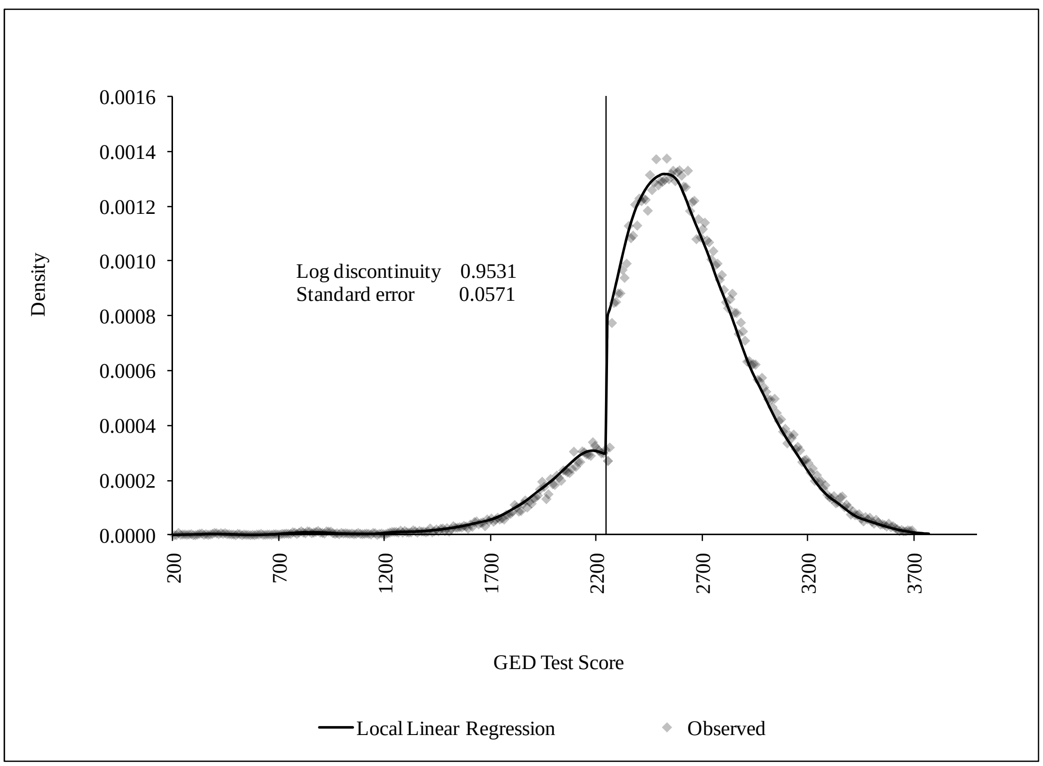


Figure 3: Distribution of First Test Score, 1995-2005

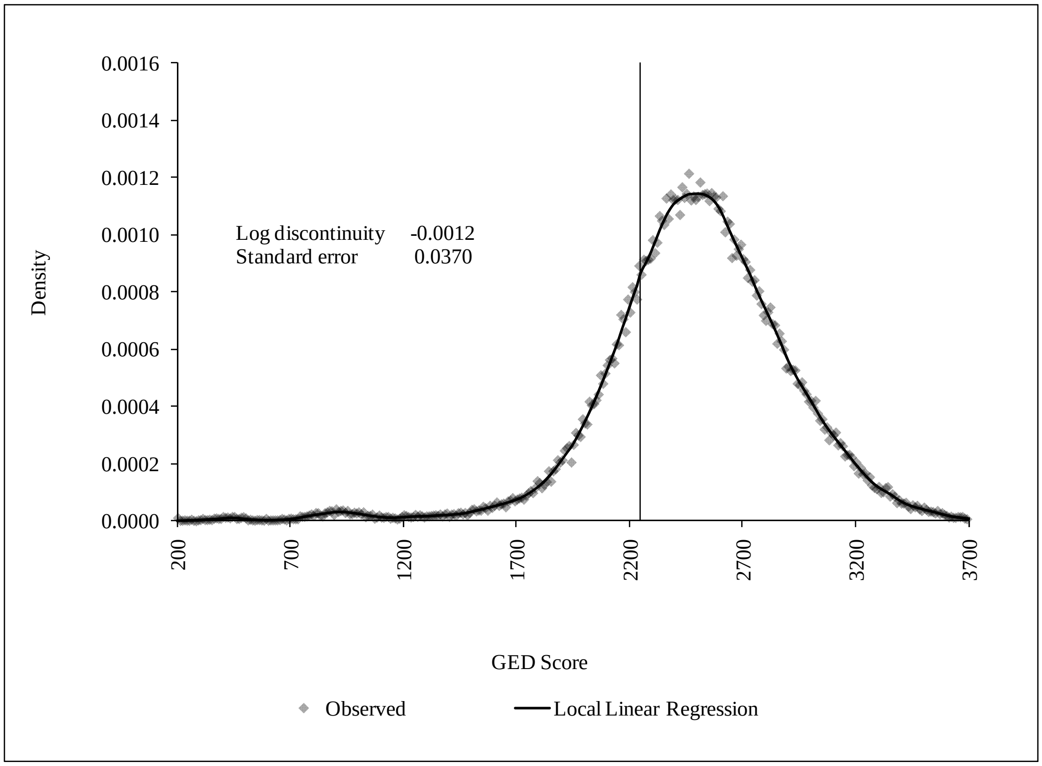


Figure 4: Regression Discontinuity Models Predicting GED and Quarterly Earnings, Men

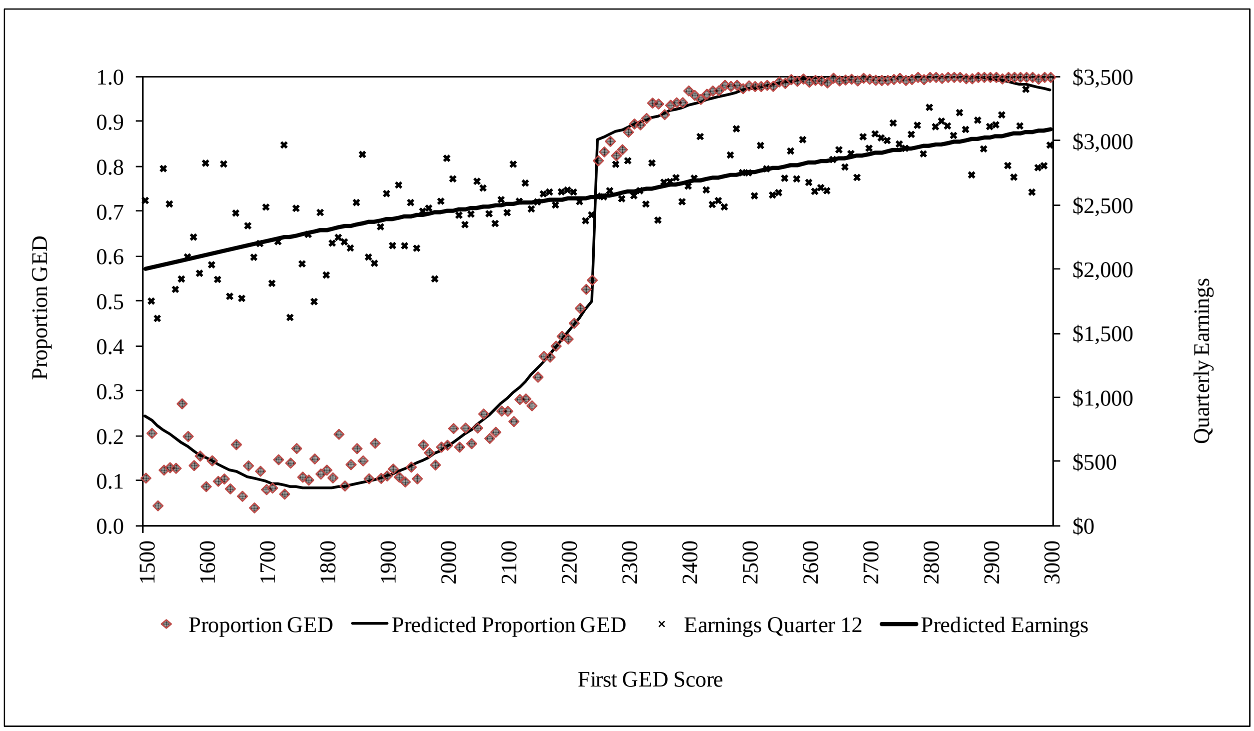


Figure 5: Estimated Multidimensional Reduced FRD and SRD GED Impacts for Men

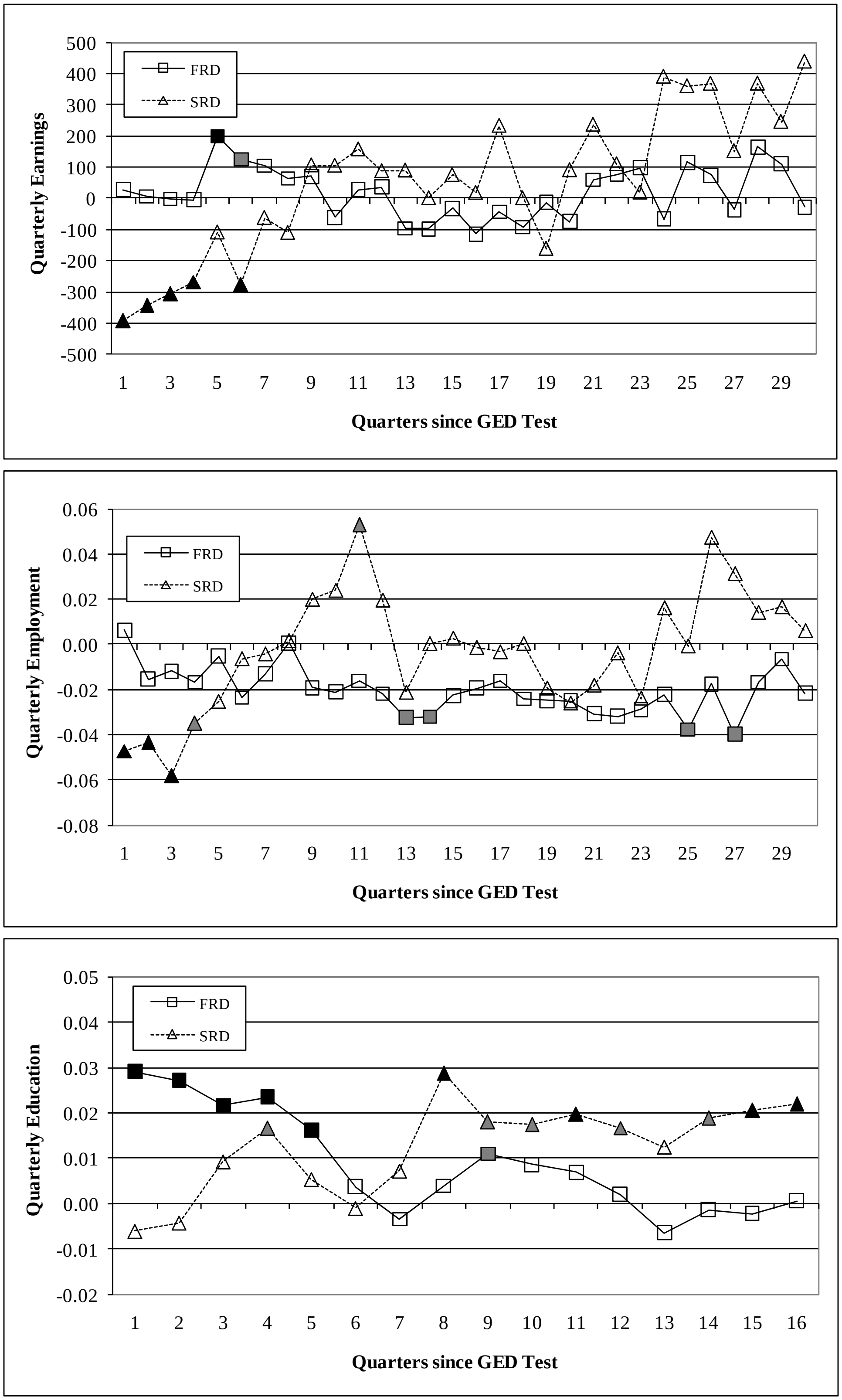
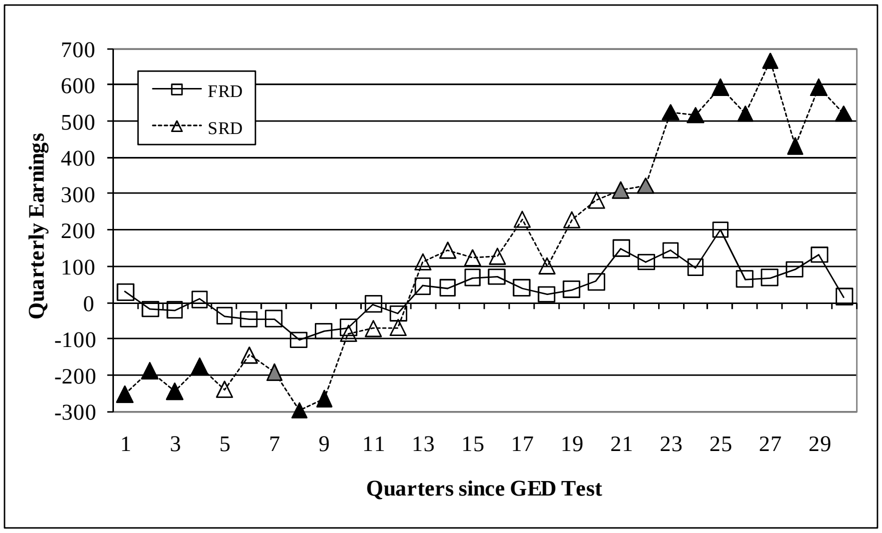
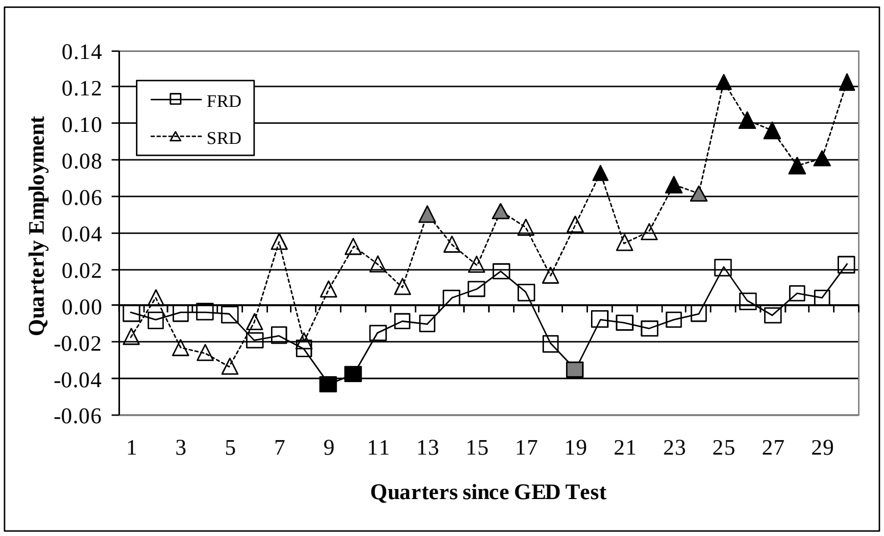
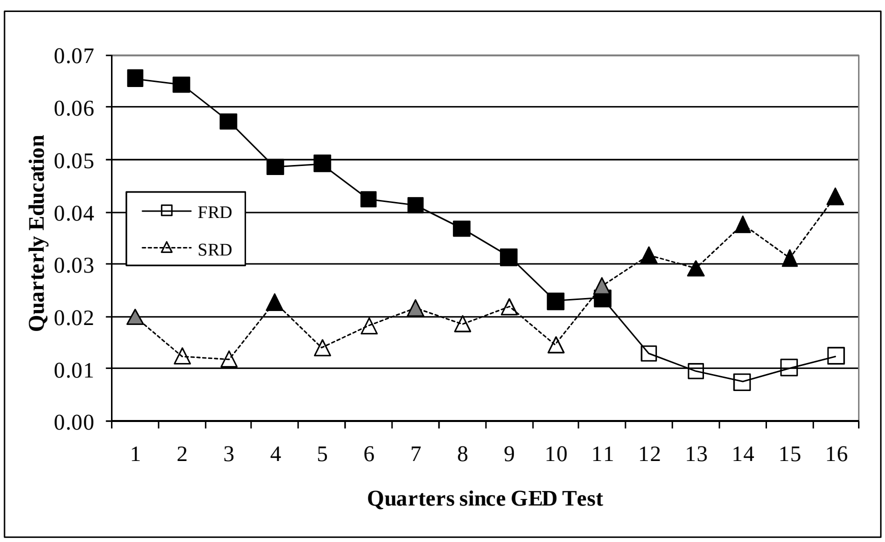


Figure 6: Estimated Multidimensional Reduced FRD and SRD GED Impacts for Women







Appendix Table A1: Descriptive Statistics

	Men	Women
Demographics		
Year first GED test		
1995-2000	70.9%	75.5%
2001	12.7%	13.1%
2002-2005	22.3%	19.4%
Nonwhite	21.6%	19.9%
GED Certification	80.4%	81.6%
Observations	44,378	41,967

Outcomes

Quarters since	Men: Earnings		Employ.	Educ.	Women: Earnings		Employ.	Educ.
1st GED test	Mean	Std. Dev.	Pct	Pct	Mean	Std. Dev.	Pct	Pct
1	$2,048	$2,862	62.1%	7.7%	$1,725	$2,237	63.1%	12.0%
2	$2,190	$2,980	62.6%	7.2%	$1,864	$2,345	64.1%	11.7%
3	$2,255	$3,043	62.1%	6.7%	$1,921	$2,639	64.2%	10.9%
4	$2,333	$3,131	61.7%	6.4%	$1,996	$2,484	64.6%	10.4%
5	$2,404	$3,173	61.7%	6.0%	$2,058	$2,537	64.4%	9.7%
6	$2,470	$3,251	61.1%	5.5%	$2,109	$2,550	64.3%	9.2%
7	$2,507	$3,352	60.7%	5.2%	$2,155	$2,611	64.0%	8.9%
8	$2,565	$3,374	60.5%	5.0%	$2,227	$2,733	64.0%	8.4%
9	$2,607	$3,407	60.0%	4.8%	$2,261	$2,714	63.9%	8.0%
10	$2,657	$3,515	59.4%	4.5%	$2,300	$2,790	63.5%	7.7%
11	$2,661	$3,574	58.7%	4.2%	$2,312	$2,799	62.7%	7.4%
12	$2,723	$3,625	58.3%	4.1%	$2,359	$2,913	62.5%	7.0%
13	$2,766	$3,682	57.5%	3.8%	$2,382	$2,895	62.2%	6.6%
14	$2,781	$3,728	56.9%	3.6%	$2,398	$2,899	61.9%	6.4%
15	$2,795	$3,906	56.4%	3.5%	$2,413	$2,948	61.3%	6.1%
16	$2,835	$3,830	56.1%	3.3%	$2,434	$2,974	60.8%	5.9%
17	$2,858	$3,850	55.6%		$2,457	$3,152	60.5%	
18	$2,888	$3,905	55.5%		$2,483	$3,056	60.2%	
19	$2,904	$3,936	55.2%		$2,473	$3,054	59.6%	
20	$2,933	$3,995	54.6%		$2,477	$3,081	59.3%	
21	$2,971	$4,064	54.5%		$2,500	$3,111	58.8%	
22	$3,016	$4,138	54.5%		$2,527	$3,226	58.4%	
23	$3,018	$4,316	53.9%		$2,517	$3,172	58.0%	
24	$3,049	$4,390	53.6%		$2,555	$3,302	57.9%	
25	$3,078	$4,223	53.4%		$2,566	$3,288	57.5%	
26	$3,104	$4,264	53.3%		$2,598	$3,306	57.5%	
27	$3,090	$4,428	52.8%		$2,576	$3,279	57.0%	
28	$3,135	$4,341	52.7%		$2,604	$3,336	56.7%	
29	$3,178	$4,374	52.7%		$2,604	$3,359	56.4%	
30	$3,211	$4,426	52.5%		$2,633	$3,392	56.2%	

Note: GED certification is measured as having ever received GED certification.