Labor-Market Returns to the GED Using Regression Discontinuity Analysis

November 9, 2010

Christopher Jepsen*
University of Kentucky
335 BA Gatton B&E Bldg
Lexington, KY 40506-0034
(859) 257-1386
jepsen@uky.edu

Peter Mueser
University of Missouri-Columbia
118 Professional Building
Columbia, MO 65211
(573) 882-6427
mueserp@missouri.edu

Kenneth Troske
University of Kentucky
335 BA Gatton B&E Bldg
Lexington, KY 40506-0034
(859) 257-1282
ktroske@uky.edu

Abstract:

In this paper, we evaluate the labor-market returns to General Educational Development (GED) certification using Missouri 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 Codes: C26, H43, I21

Keywords: Regression discontinuity, Program evaluation, GED

---

* Corresponding author. We thank Kyung-Seong Jeon and Laura Coogan for excellent research assistance, and we thank the Spencer Foundation for financial assistance. 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. 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 University of Missouri, Indiana University Purdue University Indianapolis, 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 Missouri 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, including the original application of regression discontinuity in Thislewaite and Campbell (1960). 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 effect of GED certification on either employment or earnings is not statistically significant, although our standard errors are sufficiently large that we cannot rule out sizable positive or negative effects. 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 vary from results when we estimate a sharp regression discontinuity (SRD), ignoring the ability of students to retake the test, as in Tyler (2004).

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 they have 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. When this condition is violated, it suggests that the score may be manipulated in ways that bias estimates of impact. In our context, this 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 initial attempt.

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. Examples of applications for this technique include several areas in education such as scholarships, remedial education, and financial aid, as well as areas such as job training programs and health economics.

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. Martorell and McFarlin (2008) use the initial remedial education scores in a FRD analysis of college remediation in Texas.

II.B. GED Literature

Early work on the GED used 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 using the panel nature of the NLSY data, comparing male GED recipients to male 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 10th 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.

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 results. Lofstrom and Tyler (2007) use administrative data for Texas men, but they find no signaling value of the GED—identified through the state's 1997 increase in the passing standard—on earnings, possibly due to the low GED threshold that existed prior to the 1997 change. Both papers use several techniques including regression discontinuity analysis. However, neither paper provides a meaningful test of whether the regression discontinuity analysis method used is valid, and validity is a particular concern because of students' ability to retake the test score if they do not pass. Lofstrom and Tyler (2007) conduct a robustness test where they limit the sample to students who take the test only once, but we show later that the sample of single test takers does not produce a valid regression discontinuity estimate in our Missouri data.

Heckman, Humphries, and Mader (2010) summarize the academic literature on the GED, and they conclude that the test has few if any benefits in terms of labor-market outcomes. They also point out strong non-cognitive differences between GED recipients and traditional high school graduates.

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 each individual's first GED attempt 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 are based solely on NLSY, HSB, and grouped SSA data. As discussed previously, these survey data are based on small samples of dropouts and GED recipients, and the SSA data are based on group means rather than individual regression results.

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. 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

[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 have been minor, especially since 2002.

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 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. 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 the test more than ten 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.[2] 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

[2] 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.

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.

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 from the first quarter of 1994 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 focus on 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 clearly 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, over 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 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. The final test score is therefore 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. Justification for this approach rests on the observation that only about one in seven test takers retakes the test, and that 90-95 percent of those whose overall test scores exceeds the threshold also pass the subtest minimum.

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 somewhat 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).[3]

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 has caused a dramatic redistribution in the final score.

Given that the final test score displays a marked discontinuity, it is 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

[3] These methods correspond to those recommended by McCrary (2008). Note, the kernel fitted in Figures 1-3 allows for a discontinuity at 2250.

coefficients that capture the relationship between the GED score and the outcome variable, and the coefficient α_{xlr} identifies the extent of any discontinuity.

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. There are several statistically significant differences. Males are slightly overrepresented 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 below the cut point. 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 simple alternative is to limit consideration to the cases in which individuals have not taken the test a second time. As noted above, Lofstrom and Tyler (2007) limit their sample in this way as a robustness check for their signaling model, 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 in the period 1995-2005 and did not take the test a second time through 2008. The most notable observation is that a marked discontinuity is present just as in the final test score—in fact, measured in log form, the discontinuity size is slightly larger for single test takers. 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.

The analysis here will use the first test score—for all test takers over the period 1995-2005—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.

IV. Applying Fuzzy Regression Discontinuity Methods[4]

The diagnostics suggest that the first GED test score 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 our context, the equation predicting GED certification is written:

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

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). 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. 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. The estimated parameter α_{wlr} indicates the discontinuity at the threshold.

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

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

---

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

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) $$\tau = \alpha_{ylr} / \alpha_{wlr} .$$

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) $$Y = \alpha_0 + \tau \widehat{GED} + \sum_{j=1}^{p} \beta_{lj} [D_l (T - 2250)]^j + \sum_{j=1}^{p} \beta_{rj} [D_r (T - 2250)]^j + X\eta + \upsilon ,$$

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.

As stated 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 for 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. 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 the first 14 quarters after the initial GED attempt.

Figure 4 provides a graph that illustrates the estimation methods underlying equations (1) and (2). Here the focus is earnings in quarter 12. The discontinuity assumed in equation (1) 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 appear to show a discontinuity at this point, suggesting that there is little impact on quarterly 12 earnings.

Table 3 presents estimates based on equation (1), the first stage of the two-stage equation, applied to quarter 12.[5] In Table 3, the dependent variable is a dichotomous variable for passing the GED test, and the model is estimated as OLS with clustered standard errors for each GED test score, as suggested by Lee and Card (2008). Note that the first-stage estimates for the three second-stage outcomes (quarterly income, employment, and postsecondary education) are

[5] The results from the first-stage equation vary from quarter to quarter because the dependent variable is GED certification in the quarter relevant for the impact estimation and because the sample size varies. For brevity, the table contains the results for quarter 12 after the initial GED test. The results from other quarters show a similar pattern.

identical because they are all based on the same sample and the same first-stage regression. The discontinuity at the threshold (denoted α_{wlr} in equation (1)) is associated with a 34 percentage point increase in the likelihood that men obtain GED certification, whereas the number for women is 29 percentage points. All the discontinuity variables are significant at the one-percent level (two-sided test).

Parameter estimates for the GED impact based on equation (4) are in Tables 4a through 6a for men and Table 4b through 6b for women. 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 model estimated in equation (4). As mentioned previously, we consider three dependent variables: quarterly earnings, employment, and postsecondary attendance. 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. Instrumental variables (IV) models are estimated, and the standard errors are clustered by initial GED test score.

In Table 4 the first two columns of each panel contain the coefficient and standard error when 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. Although the estimated coefficients vary from quarter to quarter, only 6 of the 60 coefficients reported are statistically significant from zero at the five-percent level (two-sided test). Thus, in most cases, we cannot reject the hypothesis that the GED has no effect on quarterly earnings, especially for women. For men, the GED does appear to be associated with higher quarterly earnings of \$300 to \$600 in quarters 5 through 9 following the initial GED test. The estimates have sizable 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 reproduce our earnings analysis limiting the sample to those who are not enrolled in public postsecondary education in Missouri during that quarter. The results from this sample are nearly identical to the results reported in the tables.

We also estimate effects taking the dependent variable as log earnings rather than earnings, limiting consideration to those with positive earnings in the year. These analyses change the structure of the question being asked, as effects on those with earnings are a result of both direct effects on those employed and effects due to selection into employment. In practice, the selection effects appear minimal (see employment results below), and the effects on earnings in this specification are overwhelmed by sampling error, as they are in our base analyses.

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), although again the results are not precisely estimated. As with earnings, the results for employment are not sensitive to the inclusion of individuals attending post-secondary education during the quarter.

Table 6 presents results for postsecondary enrollment in public institutions in Missouri. For men, GED certification is associated with increased postsecondary enrollment of approximately four percentage points in the first two quarters after the test and three percentage points in the third quarter. We also see a positive effect for men of as much as three percentage points in quarters 10 and 11. In other quarters, the effect is small and not statistically significant. For women, the effect is much larger, although it is less-precisely estimated. GED certification is associated with an increased likelihood of postsecondary attendance for the first six quarters after the initial GED attempt. The effect size is eight percentage points in the first quarter after the test, and it declines to four percentage points in the 6th quarter after the test. In later 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. In an effort to improve the estimation equations, the presented results are from models that control for various demographic factors. The exclusion of the demographics increases the standard errors by as much as one third in the quarters immediately following the first GED test, but the pattern of results is nearly identical to the reported results. Because the test changed in 2002 (and prior test scores were no longer accepted at that point), we fit models that allow the slope of the test score on GED certification and the dependent variable to differ by period. We also fit the full model separately for the period prior to and after the implementation of the new test in 2002, but sampling error overwhelms any differences by group. Finally, we combine men and women into a single sample. Although the combined sample has smaller standard errors, improvements in precision of estimates are not sufficient to resolve any of the substantive issues raised by the separate analyses.

Lee and Lemieux (2010) suggest using multiple methods, both parametric and nonparametric, for conducting regression discontinuity analysis. In response, we also fit 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 obtain 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 examine graphs of the data and the estimated functions, as well as examining how estimates varied with bandwidth (awkward sentence structure). We find that the appropriate bandwidth varies across cases considered. In no case are our final results based on these analyses seriously at variance with those presented in our models; nor is 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* face a discontinuity based on their overall score. 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. Papay, Murnane, and Willett (2009) use a similar approach when estimating the effects of the high school exit exam, which has multiple subjects, using a sharp RD design. Reardon and Robinson (2010) present five methods—including one that is essentially the same as the FRD design we employ—for conducting regression discontinuity analysis with multiple discontinuities.

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 can be written:

$$
\begin{aligned}
GED = {} & \alpha_{wl} + \sum_{j=1}^{p} \beta_{wllj}[D_{Tl}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{wllj}[D_{Tl}D_{Sl}(S-c)]^j \\
& \qquad + \phi_{wll}[D_{Tl}D_{Sl}(T-2250)(S-c)] \\
& + \alpha_{wrl}D_{Tr}D_{Sl} + \sum_{j=1}^{p} \beta_{wrlj}[D_{Tr}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{wrlj}[D_{Tr}D_{Sl}(S-c)]^j \\
& \qquad + \phi_{wrl}[D_{Tr}D_{Sl}(T-2250)(S-c)] \\
& + \alpha_{wlr}D_{Tl}D_{Sr} + \sum_{j=1}^{p} \beta_{wlrj}[D_{Tl}D_{Sr}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{wlrj}[D_{Tl}D_{Sr}(S-c)]^j \\
& \qquad + \phi_{wlr}[D_{Tl}D_{Sr}(T-2250)(S-c)] \\
& + \alpha_{wrr}D_{Tr}D_{Sr} + \phi_{w0}d_{S0} + X\eta_w + \varepsilon,
\end{aligned} \tag{5}
$$

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.[6] The dummy variable d_{S0} indicates that the lowest subtest score is zero.[7] 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 highest 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 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 IV model. The outcome equation is therefore written as:

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

[7] 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.

$$
\begin{aligned}
Y = \alpha_l + \tau GED + \sum_{j=1}^{p} \beta_{llj}[D_{Tl}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{llj}[D_{Tl}D_{Sl}(S-c)]^j \\
+ \phi_{ll}[D_{Tl}D_{Sl}(T-2250)(S-c)] \\
+ \sum_{j=1}^{p} \beta_{rlj}[D_{Tr}D_{Sl}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{rlj}[D_{Tr}D_{Sl}(S-c)]^j \\
+ \phi_{rl}[D_{Tr}D_{Sl}(T-2250)(S-c)] \\
+ \sum_{j=1}^{p} \beta_{lrj}[D_{Tl}D_{Sr}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{lrj}[D_{Tl}D_{Sr}(S-c)]^j \\
+ \phi_{lr}[D_{Tl}D_{Sr}(T-2250)(S-c)] \\
+ \sum_{j=1}^{p} \beta_{rrj}[D_{Tr}D_{Sr}(T-2250)]^j + \sum_{j=1}^{p} \gamma_{rrj}[D_{Tr}D_{Sr}(S-c)]^j \\
+ \phi_{rr}[D_{Tr}D_{Sr}(T-2250)(S-c)] + \phi_0 d_{S0} + X\eta + \upsilon
\end{aligned} \tag{6}
$$

Estimated coefficients are analogous to those in (5). The exceptions are $\beta_{rrj}, \gamma_{rrj}$ and ϕ_{rr} in (6), for which the analogous parameters are taken to be zero in equation (5)—reflecting the fact that all individuals with such scores receive GED certification. In (6) 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 (6) 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 (5) 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. 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 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 (5) as:

$$
\begin{aligned}
GED = {} & \alpha_{wl} + \sum_{j=1}^{p} \beta_{wl0j}[D_{Tl}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{w0lj}[D_{Sl}(S-c)]^{j} \\
& + \alpha_{wrl} D_{Tr} D_{Sl} + \sum_{j=1}^{p} \beta_{wrlj}[D_{Tr} D_{Sl}(T-2250)]^{j} \\
& + \alpha_{wlr} D_{Tl} D_{Sr} + \sum_{j=1}^{p} \gamma_{wlrj}[D_{Tl} D_{Sr}(S-c)]^{j} \\
& + \alpha_{wrr} D_{Tr} D_{Sr} + \phi_{w0} d_{S0} + X\eta_{w} + \varepsilon
\end{aligned}
\tag{7}
$$

The reduced version of equation (6) is specified analogously, except that here we also specify that $\beta_{rlj} = \beta_{rrj} = \beta_{r0j}$ and $\gamma_{rrj} = \gamma_{lrj} = \gamma_{0rj}$.

$$
\begin{aligned}
Y = {} & \alpha_{l} + \tau GED + \sum_{j=1}^{p} \beta_{l0j}[D_{Tl}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{0lj}[D_{Sl}(S-c)]^{j} \\
& + \sum_{j=1}^{p} \beta_{r0j}[D_{Tr}(T-2250)]^{j} + \sum_{j=1}^{p} \gamma_{0rj}[D_{Sr}(S-c)]^{j} \\
& + \phi_{0} d_{S0} + X\eta + \nu
\end{aligned}
\tag{8}
$$

The third through sixth columns in each panel of Tables 4 through 6 present results from the two models presented above, again estimated as an IV. As discussed previously, the first two columns contain results from the basic model. The reduced model presented in equations (5) and (6) yields standard errors that are generally smaller than those in the first model (columns (1) and (2)). The resulting impact estimates for earnings and employment yield substantive conclusions that are almost identical. The one exception is that the coefficients in the reduced multidimensional model for both men's and women's earnings more than five years after taking the GED are more often statistically significant at the ten-percent level (two-sided test). In the case of women, however, estimates in the reduced multidimensional model are appreciably larger than those in the full model, suggesting that the difference could well be due to bias imposed by the model's restrictions.

Estimates of the impact on educational enrollment (Table 7) display substantially greater precision for both multidimensional models and, as a result, are more often statistically significant than the basic FRD estimates. Since coefficients are generally not larger in the reduced model, this suggests that greater levels of statistical significance in this case probably do not reflect bias due to the model's restrictions. Consequently, the results from the multidimensional FRD suggest that the GED has a positive effect on enrollment in postsecondary education into the second year for men and into the third year for women.

Figures 5 and 6 illustrate the estimated impact of the GED graphically. Both figures are for the reduced multidimensional model. Figure 5 is for men, and Figure 6 is for women. The solid line and box marker in each graph contains the estimated effect, t, for the FRD model. The dashed line and triangle marker is for a sharp regression discontinuity (SRD) discussed in the next section. 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. Again, the graphs show that the FRD effects are near zero for earnings and employment for nearly all quarters, whereas the GED has a positive association with postsecondary education in the quarters immediately following the initial GED attempt.

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 such as Tyler (2004). Previous work looked at different time periods and states. Here, we can focus on individuals who took the GED between 1995 and 2005 in Missouri.

To the extent possible, we fit the same model as in 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." In other words, everyone who obtains a score of 2250 or higher on their last attempt receives GED certification, and everyone who scores below 2250 does 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 (9) below:

$$Y = \alpha + \beta PassGED + \gamma GEDAvg + X\eta + \upsilon \tag{9}$$

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.[8] *GEDAvg* is the average score on all GED attempts,[9] and X is a set of covariates. This model is estimated using OLS.

Figures 5 and 6 illustrate the SRD and FRD results for men and women, respectively. As mentioned previously, the FRD model is from the reduced multidimensional model. The figures show that the pattern of results differs by model and gender. For men's earnings, the SRD model

[8] 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.

[9] We use the average of the first and highest GED test score as a proxy for the average score. This measure is more meaningful than an average using other test scores because many test takers who retake the exam do not take all sections. Furthermore, the "highest" score as described in the previous footnote is the score used to determine GED certification.

has earnings impacts above $200 for men in quarters 9 through 16, whereas the FRD model has earnings impacts above $200 for selected quarters 5 through 9 and starting in quarter 23. For women, the earnings results for both models have impacts that generally increase in magnitude for approximately the first 22 quarters after the GED attempt. For the SRD, the earnings impacts remain large – usually in excess of $400 per quarter – for the rest of the sample period. In contrast, the earnings impacts in the FRD drop, so that the impacts are around zero for quarters 26 through 30.

The most noticeable differences between models are for employment, as illustrated by the middle panel of Figures 5 and 6. For both men and women, the SRD model has much larger employment impacts than the FRD model. In fact, the SRD results for women suggest that the GED is associated with increased employment probability, a conclusion that would not be reached using the FRD model.

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 insignificant and close to zero 12 quarters after initial GED attempt. In contrast, the SRD results do not appear to decline and produce statistically significant education effects in quarters 2 through 14.

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

The analysis above illustrates how the FRD design can be adapted to deal with the GED test, where the final test score—and therefore the receipt of GED certification—is subject to manipulation by test takers. We do not find consistently statistically significant effects of GED certification on earnings or employment. Receiving GED certification increases the chance that an individual enrolls in postsecondary education in the quarters immediately following the initial GED attempt. The effects are up to four percentage points for men and up to eight percentage points for women. Given that fewer than 12 percent of the population of GED test takers enroll in postsecondary institutions during any given quarter, this impact is substantial.

When we compare our FRD results to a SRD model which makes no attempt to allow for manipulation of the running variable, as estimated in Tyler (2004), we find substantial differences. Researchers clearly need to be careful to ensure that the appropriate RD model is estimated.

VIII. 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., 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 David Card. 2008. Regression discontinuity inference with specification error. *Journal of Econometrics* 142, no. 2 (February): 655-674.

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. 2007. Modeling the signaling value of the GED with an application to an exogenous passing standard increase in Texas. IZA Discussion Paper Series no. 2953, IZA, Bonn, Germany.

Martorell, Paco, and Isaac McFarlin. 2008. Help or Hindrance? The effects of college remediation on academic and labor-market outcomes. Unpublished manuscript.

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 (Oct):475-503.

Murnane, Richard J., John B. Willett, and John H. Tyler. 2000. Who benefits from obtaining a GED? Evidence 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, retention, and enrollment. 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. 2009. Extending the regression discontinuity approach to multiple assignment variables with an illustration using high school exit examinations. Working paper, Harvard University.

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

Thislewaite, 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.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2003. Who benefits from a GED? Evidence for females from High School and Beyond. *Economics of Education Review* 22:237-247.

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

Score Range[1]	Number	Distribution (%)	Retake (%)
0-990	1,110	1.0%	66.8%
1000-1490	994	0.9%	51.6%
1500-1740	1,607	1.4%	39.1%
1750-1990	5,342	4.6%	44.0%
2000-2090	4,776	4.1%	54.2%
2100-2140	3,210	2.8%	59.2%
2150-2190	4,013	3.5%	64.1%
2200-2240	4,652	4.0%	71.0%
2250-2290	5,278	4.6%	21.3%
2300-2340	5,947	5.1%	14.4%
2350-2490	19,986	17.3%	7.0%
2500-2740	30,179	26.1%	1.7%
2750-3090	21,744	18.8%	0.3%
3100-4000	6,841	5.9%	0.1%
Total	115,679	100.0%	16.1%

[1] Only test scores that a 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	Coeffiicient	-0.015	-0.029	0.016
	(Standard error)	(0.012)	(0.015)	(0.011)
	t-statistic	-1.2	-1.9	1.3
	Observations	99,022	84,835	100,783
Nonwhite	Coeffiicient	0.032	0.035	0.006
	(Standard error)	(0.010)	(0.012)	(0.008)
	t-statistic	3.3	3.0	0.9
	Observations	95,160	81,296	96,883
Age	Coeffiicient	-1.001	-1.265	-0.309
	(Standard error)	(0.217)	(0.274)	(0.189)
	t-statistic	-4.6	-4.6	-0.6
	Observations	99,113	84,912	100,878
Retake test	Coeffiicient	-0.028		-0.500
	(Standard error)	(0.008)		(0.006)
	t-statistic	-3.6		-91.5
	Observations	99,113		100,878
Prior earnings	Coeffiicient	-364.92	-364.17	-70.89
	(Standard error)	(241.20)	(311.38)	(209.51)
	t-statistic	-1.5	-1.2	-0.3
	Observations	99,113	84,912	100,878

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
Discontinuity	**0.34476**	**0.00689**	**0.29394**	**0.00832**
Linear - left	**0.00189**	**0.00005**	**0.00255**	**0.00006**
Linear - right	**0.00062**	**0.00003**	**0.00050**	**0.00003**
Quadratic - left	**0.00205**	**0.00007**	**0.00297**	**0.00010**
Quadratic - right	**-0.00062**	**0.00004**	**-0.00051**	**0.00004**
Earnings 4 quarters prior	0.00014	0.00084	**0.00268**	**0.00102**
Earnings 3 quarters prior	-0.00120	0.00087	-0.00008	0.00104
Earnings 2 quarters prior	0.00031	0.00091	-0.00005	0.00028
Earnings 1 quarter prior	-0.00007	0.00072	0.00023	0.00084
Nonwhite	-0.00523	0.00313	**-0.01518**	**0.00389**
Age	**-0.00363**	**0.00088**	**-0.00743**	**0.00098**
Age squared	**0.04143**	**0.01224**	**0.08461**	**0.01339**
Winter quarter	**0.01310**	**0.00363**	**0.01012**	**0.00441**
Spring quarter	**0.00816**	**0.00345**	0.00025	0.00410
Summer quarter	**0.00737**	**0.00362**	0.00471	0.00428
Constant	**0.58166**	**0.01543**	**0.73269**	**0.01800**
Observations	46,899		42,647	
Adjusted R-squared	0.6005		0.6093	

Notes: Standard errors are in parentheses. 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. Separate regressions are estimated by gender. Each regression also includes dichotomous variables for year of initial GED test. Earnings variables are measured relative to first GED attempt, such as four quarters prior to the initial GED test.

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

Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
0	-18.7	(115.7)	65.1	(99.8)	58.3	(74.0)
1	-25.6	(59.2)	81.8	(63.0)	28.8	(44.5)
2	95.6	(96.4)	51.6	(84.1)	42.8	(52.6)
3	64.6	(128.6)	35.2	(108.9)	-21.0	(71.3)
4	69.8	(121.6)	16.2	(110.3)	4.5	(86.6)
5	**296.8**	**(131.0)**	168.5	(134.7)	**235.4**	**(107.1)**
6	**430.8**	**(125.6)**	**274.8**	**(118.8)**	**233.0**	**(93.7)**
7	**429.4**	**(161.6)**	218.1	(143.4)	*201.9*	*(122.3)*
8	**438.2**	**(172.3)**	37.4	(159.3)	182.2	(129.5)
9	**598.7**	**(174.0)**	245.3	(154.7)	**271.0**	**(135.3)**
10	182.4	(189.2)	-189.6	(178.1)	-36.9	(140.6)
11	335.2	(211.8)	-134.2	(192.8)	-7.3	(162.1)
12	59.0	(182.5)	-64.4	(165.3)	22.6	(142.6)
13	137.1	(222.3)	-91.9	(177.7)	-54.4	(136.6)
14	129.7	(233.8)	-74.2	(188.3)	-29.0	(136.1)
15	72.3	(219.8)	11.3	(189.0)	14.1	(143.7)
16	5.6	(227.9)	124.9	(195.5)	-29.2	(162.7)
17	134.1	(222.8)	-24.1	(191.6)	-39.2	(155.5)
18	135.9	(254.3)	-4.9	(202.8)	19.8	(173.8)
19	59.0	(303.2)	-142.3	(229.9)	22.9	(188.7)
20	-24.9	(351.5)	-85.3	(253.5)	-17.1	(199.9)
21	147.1	(319.5)	-84.4	(249.5)	129.5	(197.9)
22	269.4	(341.5)	10.7	(265.3)	194.5	(211.4)
23	316.4	(320.5)	262.9	(222.2)	300.9	(194.3)
24	197.7	(305.2)	-5.3	(228.9)	39.7	(200.1)
25	**618.4**	**(311.2)**	382.7	(243.0)	*328.2*	*(189.4)*
26	351.3	(339.1)	**514.2**	**(259.5)**	*345.2*	*(206.0)*
27	54.5	(466.9)	151.8	(353.6)	26.1	(241.8)
28	460.9	(436.6)	550.4	(330.9)	*456.5*	*(248.3)*
29	203.7	(403.7)	431.1	(300.8)	311.1	(255.5)
30	59.4	(479.6)	153.2	(386.6)	91.6	(304.2)

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.

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

Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
0	-115.5	(117.7)	-4.9	(90.5)	-21.7	(69.1)
1	-59.1	(62.2)	-0.9	(52.4)	32.2	(38.4)
2	-72.5	(101.9)	6.4	(73.0)	13.3	(56.1)
3	43.5	(113.3)	58.2	(104.7)	39.7	(66.0)
4	3.1	(169.0)	33.6	(126.3)	2.5	(82.5)
5	4.9	(175.9)	-31.5	(128.2)	-26.5	(100.7)
6	120.4	(178.7)	-46.1	(128.6)	-22.3	(102.2)
7	99.3	(198.8)	-110.9	(140.6)	23.5	(108.1)
8	36.9	(190.9)	-140.1	(148.9)	-74.9	(97.8)
9	143.1	(194.0)	61.2	(156.0)	18.8	(106.3)
10	127.0	(218.6)	-30.3	(146.4)	-62.8	(118.5)
11	201.0	(236.7)	41.8	(161.2)	96.1	(117.1)
12	162.9	(238.1)	20.7	(156.6)	77.7	(123.7)
13	114.2	(232.8)	141.5	(173.4)	**256.6**	**(123.2)**
14	*334.3*	*(197.8)*	*272.2*	*(161.7)*	**357.9**	**(117.1)**
15	406.3	(249.4)	*276.9*	*(167.4)*	**327.1**	**(127.6)**
16	452.7	(287.5)	184.9	(185.9)	213.8	(156.2)
17	268.1	(300.1)	6.5	(199.6)	157.5	(165.7)
18	359.8	(322.1)	73.1	(214.8)	150.2	(183.6)
19	259.7	(354.4)	-11.3	(245.8)	178.9	(191.9)
20	354.0	(383.7)	-41.3	(265.3)	197.8	(201.7)
21	187.7	(400.3)	47.4	(275.9)	313.5	(199.7)
22	478.4	(393.1)	73.5	(290.5)	*383.4*	*(210.7)*
23	311.7	(392.7)	-19.7	(302.0)	*332.4*	*(190.6)*
24	425.9	(363.0)	164.8	(263.7)	*301.5*	*(176.3)*
25	52.9	(331.7)	-35.5	(255.8)	105.7	(169.3)
26	-99.4	(358.8)	-175.1	(266.2)	-10.4	(184.5)
27	29.2	(427.2)	-176.2	(295.0)	42.1	(201.3)
28	82.0	(393.3)	-136.2	(270.5)	25.7	(194.2)
29	-140.3	(394.8)	-321.4	(297.7)	38.0	(212.3)
30	-301.4	(380.1)	-323.9	(276.6)	-59.6	(210.1)

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.

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

Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
0	0.013	(0.035)	-0.011	(0.031)	-0.010	(0.020)
1	0.015	(0.015)	0.015	(0.016)	0.007	(0.011)
2	-0.004	(0.019)	-0.011	(0.017)	-0.006	(0.011)
3	*0.037*	*(0.021)*	-0.007	(0.018)	-0.015	(0.013)
4	0.010	(0.027)	-0.014	(0.019)	-0.008	(0.014)
5	0.028	(0.025)	-0.011	(0.021)	0.003	(0.017)
6	0.013	(0.022)	0.000	(0.020)	-0.008	(0.016)
7	0.007	(0.026)	-0.018	(0.020)	-0.009	(0.016)
8	0.026	(0.026)	-0.009	(0.022)	0.004	(0.017)
9	0.024	(0.034)	-0.025	(0.026)	-0.025	(0.022)
10	0.016	(0.034)	-0.031	(0.026)	-0.022	(0.023)
11	0.007	(0.034)	**-0.056**	**(0.028)**	-0.025	(0.026)
12	-0.022	(0.028)	-0.032	(0.025)	-0.030	(0.022)
13	-0.035	(0.033)	-0.047	(0.029)	-0.037	(0.024)
14	0.026	(0.032)	-0.009	(0.026)	-0.014	(0.021)
15	0.036	(0.030)	-0.023	(0.027)	-0.019	(0.021)
16	0.003	(0.037)	-0.015	(0.034)	-0.014	(0.028)
17	0.003	(0.030)	-0.009	(0.032)	-0.010	(0.024)
18	0.029	(0.042)	-0.012	(0.031)	-0.013	(0.027)
19	0.013	(0.039)	-0.031	(0.034)	-0.016	(0.027)
20	0.006	(0.044)	-0.002	(0.033)	-0.012	(0.027)
21	-0.027	(0.042)	-0.037	(0.039)	-0.034	(0.031)
22	-0.009	(0.040)	-0.045	(0.036)	-0.016	(0.030)
23	-0.001	(0.043)	0.004	(0.036)	0.006	(0.029)
24	-0.020	(0.041)	-0.019	(0.033)	-0.002	(0.026)
25	0.052	(0.037)	0.027	(0.036)	0.000	(0.027)
26	0.009	(0.043)	0.040	(0.035)	0.025	(0.026)
27	0.011	(0.044)	0.013	(0.036)	-0.015	(0.027)
28	0.045	(0.053)	0.053	(0.042)	0.022	(0.030)
29	0.017	(0.060)	0.026	(0.046)	0.016	(0.037)
30	-0.031	(0.062)	-0.009	(0.050)	-0.011	(0.037)

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.

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

Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
0	-0.020	(0.040)	0.006	(0.030)	0.012	(0.025)
1	-0.009	(0.016)	-0.002	(0.017)	-0.005	(0.010)
2	-0.013	(0.022)	-0.002	(0.017)	-0.001	(0.013)
3	-0.011	(0.023)	-0.008	(0.023)	-0.002	(0.015)
4	-0.015	(0.028)	0.007	(0.022)	0.002	(0.015)
5	0.010	(0.028)	0.025	(0.027)	0.006	(0.018)
6	0.003	(0.034)	-0.004	(0.025)	-0.003	(0.018)
7	0.039	(0.034)	0.010	(0.026)	0.005	(0.019)
8	0.017	(0.037)	-0.001	(0.031)	-0.003	(0.021)
9	0.009	(0.040)	-0.003	(0.031)	-0.011	(0.020)
10	0.002	(0.032)	-0.019	(0.026)	-0.027	(0.019)
11	-0.014	(0.034)	-0.001	(0.030)	-0.003	(0.018)
12	0.020	(0.052)	0.026	(0.031)	0.020	(0.024)
13	-0.018	(0.048)	-0.002	(0.033)	0.004	(0.024)
14	0.070	(0.043)	*0.052*	*(0.030)*	**0.051**	**(0.022)**
15	0.089	(0.056)	0.048	(0.035)	0.039	(0.027)
16	0.049	(0.049)	0.013	(0.034)	0.017	(0.027)
17	0.027	(0.053)	0.017	(0.030)	0.002	(0.026)
18	0.040	(0.046)	0.007	(0.038)	0.002	(0.029)
19	0.023	(0.059)	-0.012	(0.040)	-0.007	(0.030)
20	0.064	(0.053)	-0.007	(0.038)	0.010	(0.029)
21	0.008	(0.066)	-0.024	(0.046)	0.015	(0.035)
22	0.008	(0.055)	-0.057	(0.049)	-0.014	(0.035)
23	-0.007	(0.077)	-0.055	(0.049)	-0.012	(0.036)
24	0.015	(0.061)	-0.010	(0.042)	-0.003	(0.031)
25	0.017	(0.059)	-0.001	(0.040)	0.019	(0.032)
26	0.015	(0.048)	0.003	(0.035)	-0.004	(0.028)
27	0.007	(0.049)	-0.015	(0.039)	0.006	(0.027)
28	-0.017	(0.039)	-0.010	(0.037)	-0.005	(0.026)
29	-0.036	(0.057)	-0.037	(0.039)	-0.010	(0.029)
30	-0.003	(0.059)	-0.019	(0.043)	0.003	(0.032)

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.

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

Men						
Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
1	**0.040**	**(0.006)**	**0.038**	**(0.006)**	**0.029**	**(0.004)**
2	**0.044**	**(0.010)**	**0.044**	**(0.008)**	**0.030**	**(0.006)**
3	**0.030**	**(0.009)**	**0.044**	**(0.008)**	**0.022**	**(0.007)**
4	0.021	(0.011)	**0.037**	**(0.009)**	**0.019**	**(0.008)**
5	0.016	(0.012)	**0.029**	**(0.011)**	0.012	(0.010)
6	0.012	(0.011)	*0.020*	*(0.012)*	0.001	(0.009)
7	-0.003	(0.011)	0.002	(0.010)	-0.001	(0.008)
8	-0.001	(0.012)	-0.003	(0.010)	0.003	(0.008)
9	0.019	(0.014)	0.011	(0.011)	0.009	(0.008)
10	**0.030**	**(0.013)**	0.013	(0.011)	0.009	(0.008)
11	**0.019**	**(0.009)**	0.007	(0.009)	0.005	(0.008)
12	0.005	(0.011)	0.001	(0.010)	-0.002	(0.008)
13	-0.003	(0.012)	-0.007	(0.009)	-0.011	(0.008)
14	0.009	(0.013)	0.002	(0.010)	-0.001	(0.007)

Women						
Quarters since 1st GED test	Basic		Multidimensional Full model		Multidimensional Reduced model	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
1	**0.081**	**(0.012)**	**0.078**	**(0.012)**	**0.066**	**(0.008)**
2	**0.067**	**(0.012)**	**0.065**	**(0.012)**	**0.056**	**(0.008)**
3	**0.042**	**(0.012)**	**0.071**	**(0.015)**	**0.055**	**(0.009)**
4	**0.058**	**(0.014)**	**0.064**	**(0.015)**	**0.047**	**(0.011)**
5	**0.053**	**(0.016)**	**0.062**	**(0.014)**	**0.045**	**(0.012)**
6	**0.041**	**(0.020)**	**0.054**	**(0.017)**	**0.042**	**(0.014)**
7	*0.038*	*(0.022)*	**0.034**	**(0.015)**	**0.030**	**(0.012)**
8	0.030	(0.024)	0.026	(0.017)	**0.026**	**(0.013)**
9	0.033	(0.021)	**0.036**	**(0.018)**	**0.032**	**(0.012)**
10	0.014	(0.016)	0.028	(0.020)	*0.023*	*(0.013)*
11	0.020	(0.017)	*0.031*	*(0.017)*	**0.022**	**(0.011)**
12	-0.002	(0.017)	0.006	(0.016)	0.002	(0.010)
13	-0.009	(0.017)	0.004	(0.019)	0.004	(0.011)
14	-0.006	(0.019)	-0.007	(0.019)	0.005	(0.011)

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.

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

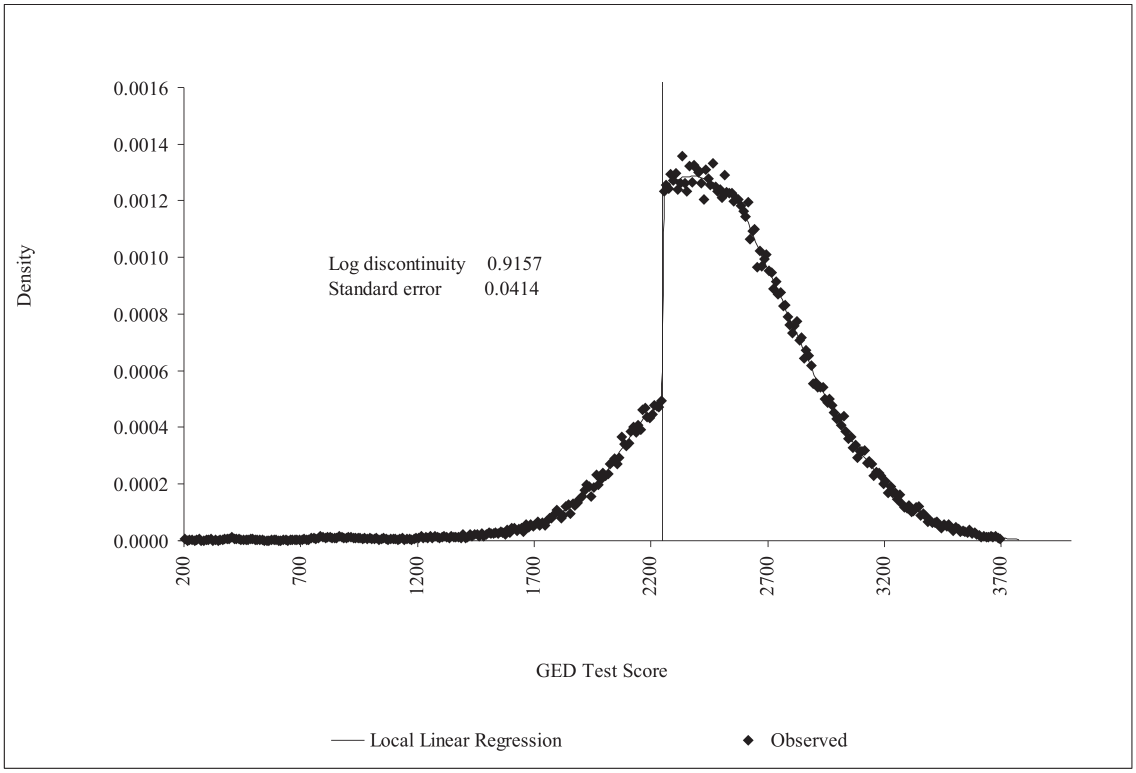


Figure 2: Distribution of First Test Score for Single Test Takers, 1995-2005

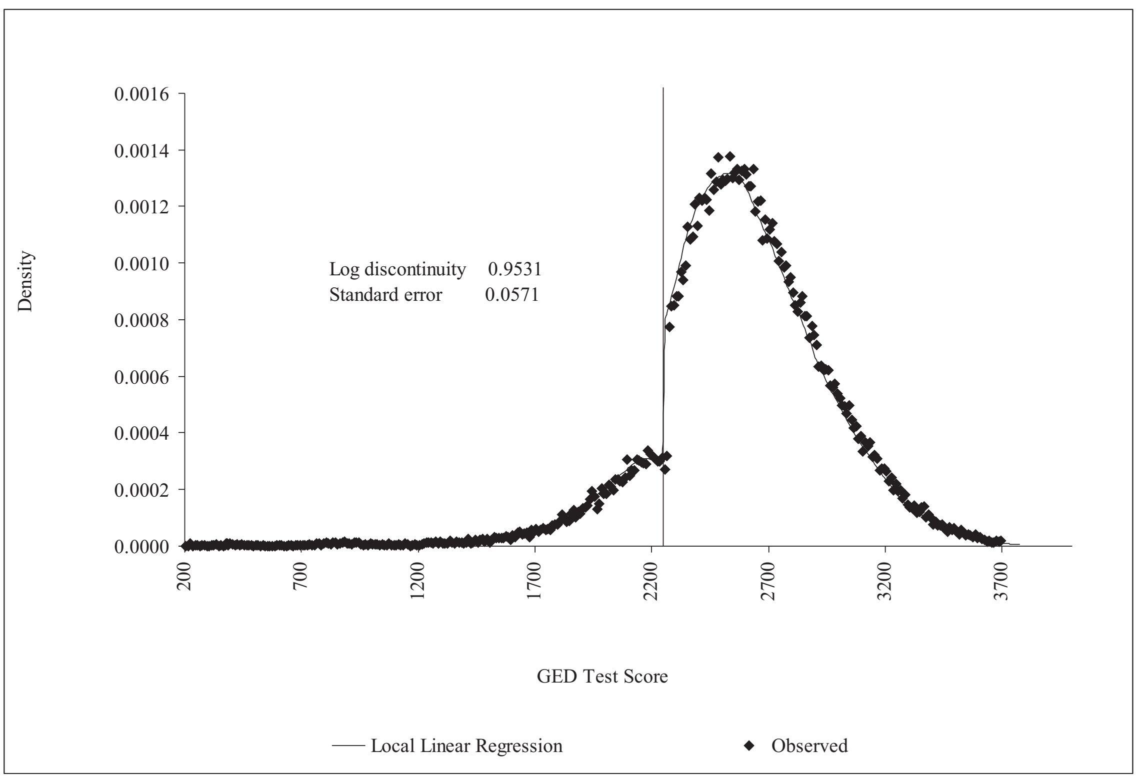


Figure 3: Distribution of First Test Score, 1995-2005

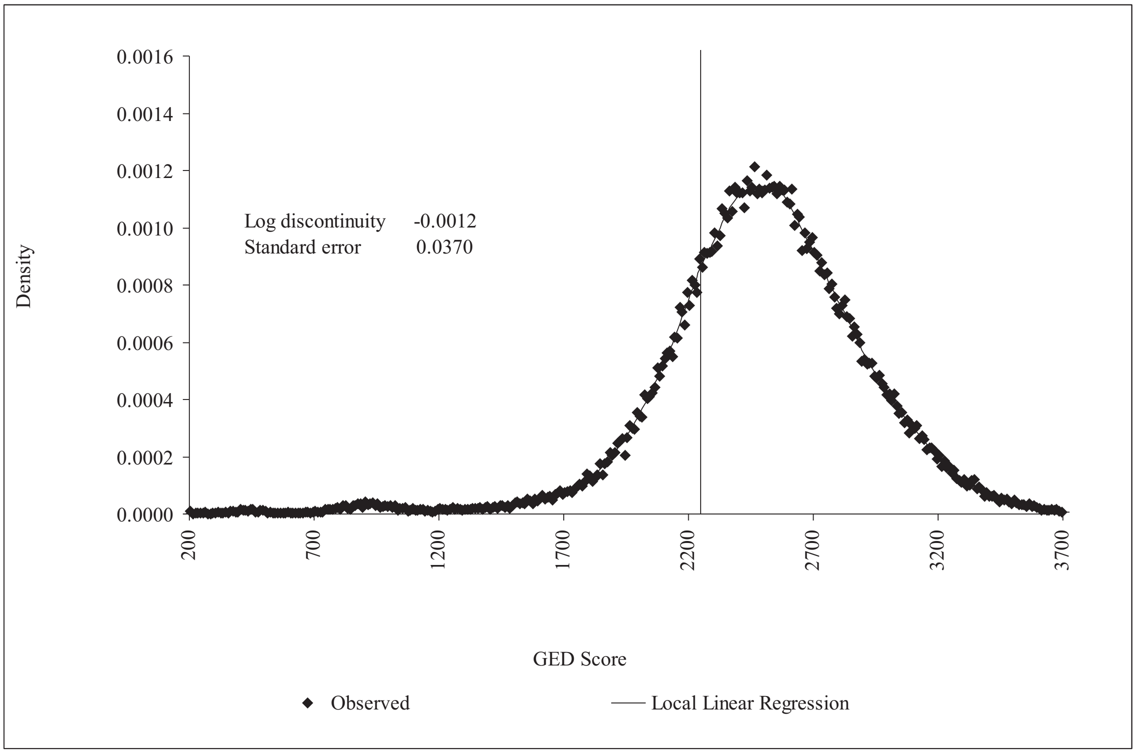


Figure 4: Regression Discontinuity Models Predicting GED and Quarterly Earnings, Men

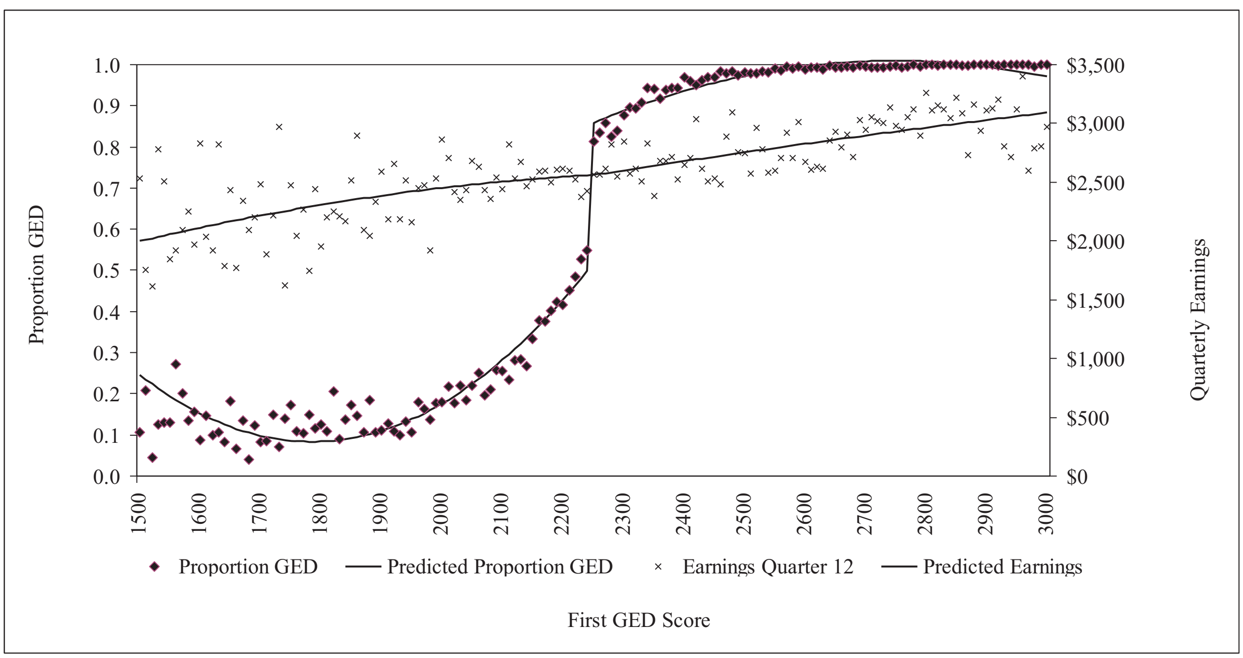


Figure 5: Estimated Multidimensional Reduced FRD and SRD GED Impacts for Men

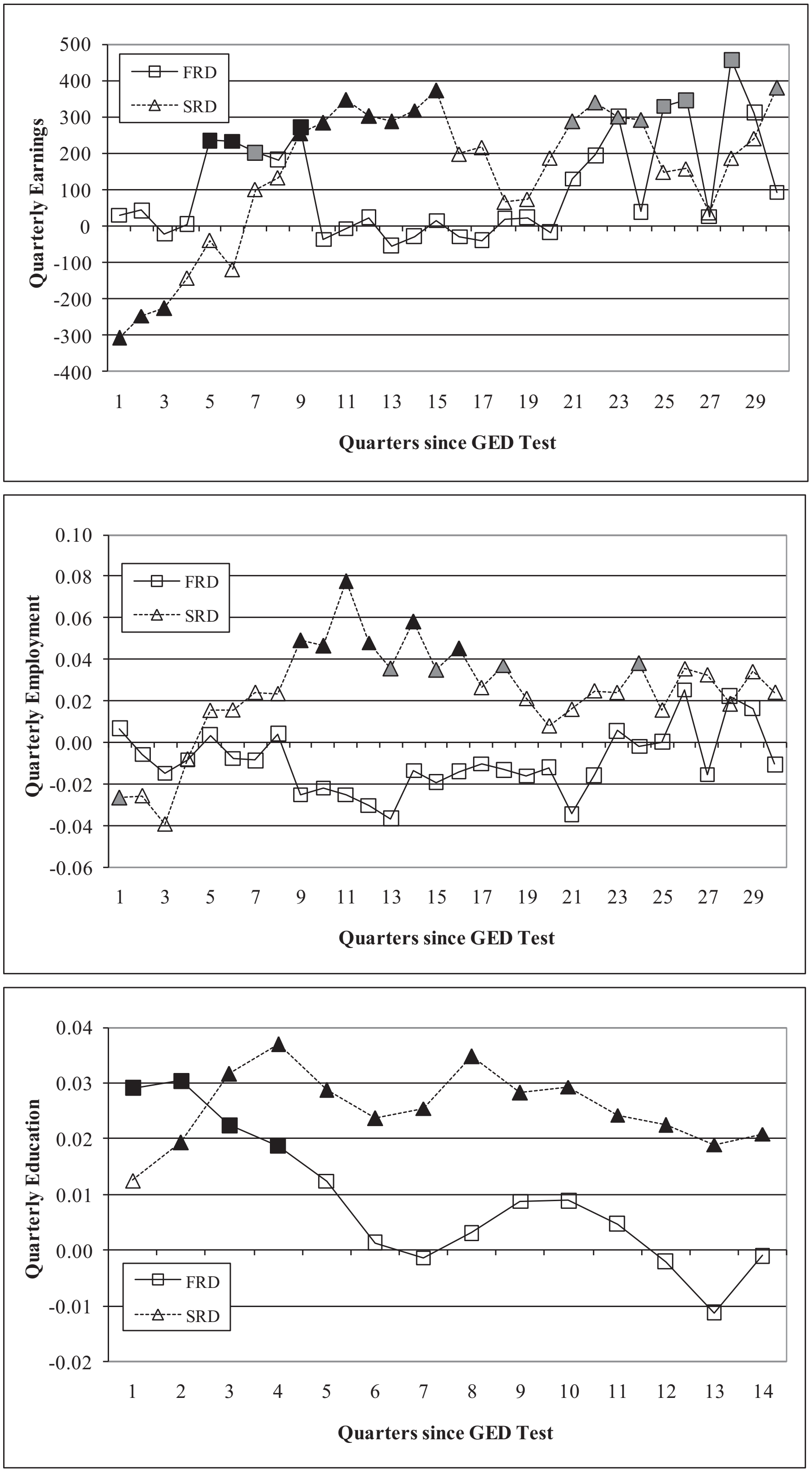
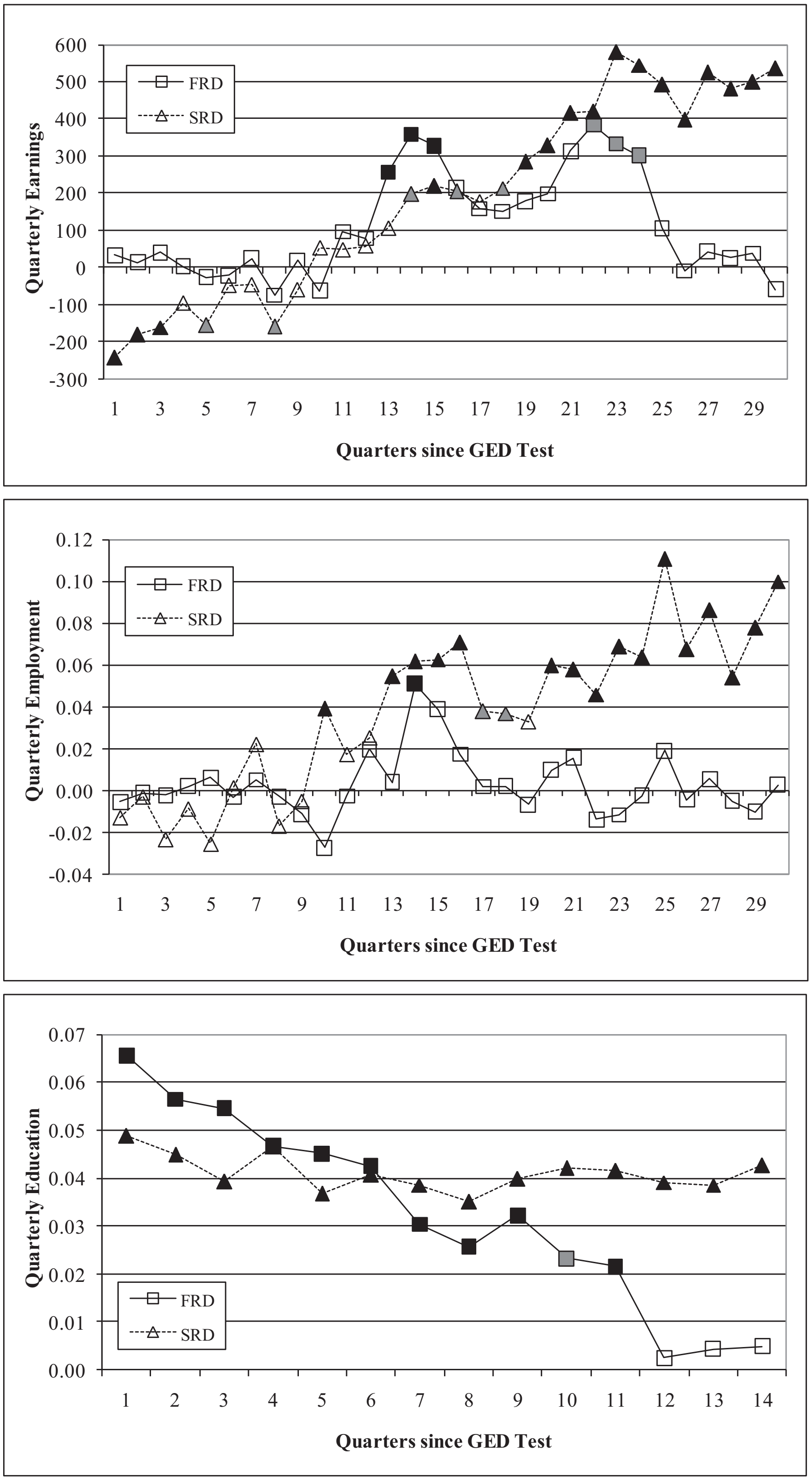


Figure 6: Estimated Multidimensional Reduced FRD and SRD GED Impacts for Women



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	80.7%				81.7%			
Outcomes								
Quarters since 1st GED test	Earnings Mean	Earnings Std. Dev.	Employ. Pct	Educ. Pct	Earnings Mean	Earnings Std. Dev.	Employ. Pct	Educ. Pct
1	$2,045	$2,881	61.5%	7.4%	$1,731	$2,257	62.9%	11.9%
2	$2,192	$3,008	62.0%	6.9%	$1,870	$2,363	64.0%	11.5%
3	$2,255	$3,072	61.4%	6.4%	$1,928	$2,655	64.0%	10.8%
4	$2,332	$3,157	61.0%	6.0%	$2,003	$2,506	64.5%	10.2%
5	$2,403	$3,200	60.9%	5.7%	$2,064	$2,553	64.2%	9.6%
6	$2,470	$3,279	60.4%	5.3%	$2,115	$2,568	64.2%	9.1%
7	$2,506	$3,374	60.0%	5.0%	$2,160	$2,631	63.8%	8.8%
8	$2,564	$3,402	59.7%	4.8%	$2,231	$2,747	63.8%	8.3%
9	$2,605	$3,435	59.2%	4.6%	$2,265	$2,730	63.7%	7.9%
10	$2,657	$3,548	58.6%	4.3%	$2,304	$2,804	63.3%	7.6%
11	$2,659	$3,598	58.0%	4.0%	$2,316	$2,814	62.5%	7.2%
12	$2,722	$3,651	57.6%	3.9%	$2,362	$2,923	62.4%	6.9%
13	$2,764	$3,709	56.8%	3.6%	$2,388	$2,913	62.0%	6.5%
14	$2,782	$3,756	56.3%	3.4%	$2,402	$2,910	61.7%	6.3%
15	$2,805	$3,931	55.9%		$2,421	$2,962	61.2%	
16	$2,852	$3,849	55.9%		$2,447	$2,984	60.9%	
17	$2,903	$3,895	55.6%		$2,482	$3,169	60.8%	
18	$2,952	$3,975	55.7%		$2,515	$3,066	60.6%	
19	$2,970	$4,014	55.3%		$2,507	$3,087	60.0%	
20	$2,996	$4,055	54.9%		$2,515	$3,117	59.7%	
21	$3,033	$4,246	54.7%		$2,537	$3,143	59.2%	
22	$3,074	$4,330	54.6%		$2,559	$3,259	58.7%	
23	$3,068	$4,382	53.9%		$2,548	$3,195	58.4%	
24	$3,097	$4,440	53.6%		$2,585	$3,334	58.1%	
25	$3,119	$4,269	53.4%		$2,589	$3,306	57.6%	
26	$3,157	$4,304	53.3%		$2,623	$3,330	57.7%	
27	$3,154	$4,497	52.9%		$2,598	$3,299	57.2%	
28	$3,211	$4,405	52.9%		$2,625	$3,353	56.7%	
29	$3,248	$4,439	52.8%		$2,623	$3,375	56.6%	
30	$3,262	$4,480	52.6%		$2,650	$3,409	56.3%	

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