5-4-2016

Labor Market Returns to the GED Using Regression Discontinuity Analysis

Christopher Jepsen
University College Dublin, Ireland

Peter Mueser
University of Missouri

Kenneth R. Troske
University of Kentucky, ktroske@uky.edu

Labor Market Returns to the GED Using Regression Discontinuity Analysis

Notes/Citation Information
Published in *Journal of Political Economy*, v. 124, no. 3, p. 621-649.

Digital Object Identifier (DOI)
https://doi.org/10.1086/686245

Labor Market Returns to the GED Using Regression Discontinuity Analysis

Christopher Jepsen

University College Dublin and Institute for the Study of Labor (IZA)

Peter Mueser

University of Missouri, Columbia and Institute for the Study of Labor (IZA)

Kenneth Troske

University of Kentucky and Institute for the Study of Labor (IZA)

We evaluate returns to General Educational Development (GED) certification for high school dropouts using state administrative data. We apply a fuzzy regression discontinuity method to account for test takers retaking the test. For women we find that GED certification has no statistically significant effect on either employment or earnings. For men we find a significant increase in earnings in the second year after taking the test but no impact in subsequent years. GED certification increases postsecondary school enrollment by 4–8 percentage points. Our results differ from regression discontinuity approaches that fail to account for test retaking.

I. Introduction

Labor market opportunities for high school dropouts have declined substantially in recent years. Certification on the General Educational De-

We thank Kyung-Seong Jeon and Laura Coogan for excellent research assistance, and we thank the Spencer Foundation for financial assistance. We thank the anonymous referees, Tom Ahn, Sandra Black, James Heckman, Paco Martorell, Richard Murnane, John Papay,

Electronically published May 4, 2016
[*Journal of Political Economy*, 2016, vol. 124, no. 3]

velopment (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 using state administrative data. We apply a fuzzy regression discontinuity method to account for the fact that GED test takers can repeatedly retake the test until they pass it and that GED certification depends on meeting subtest requirements. Previous analyses of the GED based on regression discontinuity methods have used as a forcing variable a composite test score that includes scores obtained by retaking the test. Because this approach ignores the fact that individuals choose whether to retake the test, estimates are subject to bias. Our approach, based on the discontinuity on the score from the first test taken, can be applied to other situations in which program participation is determined by a score on a test that can be retaken multiple times. Examples of tests to which this technique can be applied include civil service exams, bar exams, votes for unionization, and licensure exams such as drivers' licenses.

We find that, for dropouts who take the GED and score near the cutoff, the estimated effects of GED certification on either employment or earnings are generally small and not statistically significant, with the exception of an increase in earnings for males five to nine quarters after taking the test. GED certification increases postsecondary participation in the year following certification by up to 4 percentage points for men and up to 8 percentage points for women. Finally, the results from our preferred model often differ from results of a sharp regression discontinuity design, which ignores the ability of students to retake the test.

Jeffrey Smith, and seminar participants at the 2009 Association for Public Policy and Management conference, the 2010 Midwest Economics Association conference, the 2010 Society of Labor Economists/European Association of Labour Economists 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 2012 Institute for Research on Poverty Summer Research Workshop, the 12th Institute for the Study of Labor/Society of Labor Economists Transatlantic Meeting of Labor Economists, University of Missouri, Indiana University–Purdue University Indianapolis, MDRC, and Ohio State University for useful comments. All opinions and errors are the authors' sole responsibility. The data used in this study are proprietary data from Missouri's Department of Elementary and Secondary Education (DESE) and Missouri's Department of Labor and Industrial Relations (DOLIR), and we are precluded from sharing the data with others. Researchers interested in replicating our results would need to apply to DESE and DOLIR for the data. We would be happy to provide assistance and computer code to go from the raw data to the results of the paper.

II. Relation to Previous Literature

Heckman, Humphries, and Kautz (2014) and Heckman and Kautz (2014) provide the most comprehensive analysis of the labor market returns to the GED. In addition to reviewing previous work, starting with Cameron and Heckman (1993), they estimate the GED impact across six survey data sets, including the National Longitudinal Survey of Youth and the National Educational Longitudinal Survey of 1988. In general, they find no evidence of a GED impact on labor market earnings or participation for men compared to dropouts, but for women they find some evidence of higher annual earnings, driven by higher labor force participation. They show that this pattern of results is largely consistent with much of the previous work using survey data.[1]

Among the most widely read papers related to the GED is Tyler, Murnane, and Willett (2000b), which reports positive effects of GED certification on earnings for whites (males and females) but not for nonwhites. Estimates of the effects of GED are based on a comparison of test takers in states with different GED passing thresholds, essentially comparing test takers who receive certification with others who have identical test scores but do not. Heckman et al. (2014) point out that this method is similar to a regression discontinuity approach, as it focuses on differences in GED certification between individuals with similar GED test scores.[2] Using a similar approach for a smaller set of states, Tyler, Murnane, and Willett (2000a) find a consistent, positive association between GED certification and annual earnings for nonwhite males, white females, and nonwhite females.

A major limitation in the studies by Tyler et al. (2000a, 2000b) is that they use a composite test score based on multiple test attempts as their forcing variable in a regression discontinuity analysis. Rubinstein (2003) and Heckman et al. (2014) highlight multiple sources of bias in the estimated GED effect using this method, as well as pointing out concerns in the disparity in findings between the two Tyler et al. papers. One source of potential bias is that test takers can manipulate the composite score by retaking parts of the test until they receive a passing score, which potentially invalidates the basic assumption of regression discontinuity analysis. As we show below, the first GED test attempt generates a valid regression discontinuity estimator, but Tyler et al.'s composite score does not.

[1] Because of the extensive literature review contained in the work of Heckman and coauthors, we focus our discussion on papers that use a methodology similar to ours for estimating the impact of the GED. Readers who are interested in a more general review of this literature should refer to Heckman et al. (2014) and Heckman and Kautz (2014).

[2] The web appendix from Heckman et al. (2014), available at http://jenni.uchicago.edu/Studies_of_GED/, contains a detailed explanation of the identification strategy used in Tyler et al. (2000b), along with an explanation of the limitations of that strategy.

Papers related to those by Tyler et al. include Tyler (2004) and Loftstrom and Tyler (2008), both of which utilize individual-level administrative earnings records matched with records of GED test takers to compare male GED recipients with dropouts who took but did not pass the GED. Using data from Florida, Tyler (2004) finds positive long-run earnings effects, whereas using Texas data, Lofstrom and Tyler (2008) find no impact of the GED—identified through the state's 1997 increase in the passing standard—on earnings.

Both the Tyler (2004) and Loftstrom and Tyler (2008) papers use several techniques including ordinary least squares (OLS) and individual fixed-effects models, as well as regression discontinuity (RD) analysis. While OLS and fixed-effects models suffer from potential bias due to omitted measures of motivation or noncognitive ability, RD models offer the potential of overcoming such problems. However, these studies use the same composite test score as a forcing variable in their RD design as Tyler et al. (2000a, 2000b), so they are subject to the same biases as those studies.

Our analysis provides several contributions to the GED literature. First, we use administrative data from a single state for nearly 100,000 individuals who took the GED between 1995 and 2005. We match these data with earnings data covering the period 1993–2009, 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. The use of administrative data complements the work using panel survey data with much smaller sample sizes.

Second, as noted above, the previous GED research using RD 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 use a technique based on the score from the first GED test to produce valid RD estimates. Thus, we provide the first estimates of GED impacts based on administrative data that account for the retaking behavior of GED test takers.

Little previous work has addressed the issue of test retaking with respect to RD models. In partial exceptions, Pantal, Podgursky, and Mueser (2006) consider a test used for allocating a scholarship, and Martorell and McFarlin (2011) consider a college placement test. Both use a first test score as a forcing variable to eliminate the impact of test retaking in an RD framework, but neither considers the significance of this choice.[3]

[3] Cellini, Ferreira, and Rothstein (2010) implement an RD model examining the impact of municipal bond referenda that allows for multiple referenda, but their methods of estimation rely on assumptions that are not applicable in a testing context.

Finally, it is important to recognize that our analysis compares outcomes for people who take and pass the GED with those for individuals who take and fail to pass the GED. To the extent that high school dropouts who take the GED are more motivated than high school dropouts who do not take the GED, and this motivation is valued in the labor market, our analysis implicitly controls for possible differences in motivation. Of course, this also means that our results need not generalize to high school dropouts who choose not to take the GED.

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. The GED test is a 7½-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 current version of the GED has been in place since 2014 (GED Testing Service 2013). Subject to certain constraints, states set their own criteria for certification based on test performance, but differences between states are minor. As recently as the 1990s, there were some differences across states in the score required for certification, but such differences were small. By the 2000s, standards for passing were all but universal across states.

To obtain GED certification, test takers in our data 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 is based on a composite score computed as the sum of the highest score on each subtest taken over the prior 2 years; that is, each subtest score is "valid" for 2 years before it expires. Many individuals with scores below the required thresholds retake the test—often several times—within 2 years, and they often retake only certain subjects rather than retaking the entire exam.[4]

The 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 2002–5).

[4] Individuals can take the 2002 GED test up to six times in any 2-year period. 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.

Our basic sample consists of any individual who took the GED test for the first time in one state between 1995 and 2005.[5] For each individual taking the test within this period, we have access to data on the most recent 10 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 10 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.[6] We exclude individuals with missing information on gender or race. Individuals who received their GED through the DANTES program, which provides state certification for tests taken by military personnel outside the state, are also excluded because test scores are not reported for program participants who took the GED test through this program. Finally, we exclude individuals who took the GED as part of the GED Option program because these individuals are still enrolled in high school and therefore are fundamentally different from our sample of high school dropouts who take the GED. Descriptive statistics for the regression sample are in Appendix tables A1 and A2.[7]

Quarterly earnings in all unemployment insurance covered jobs are available as reported by employers in states' unemployment insurance programs for the state and a neighboring state. Very few of the state's residents commute to states other than these two. We use data through the second quarter of 2009.

We also look at how passing the GED affects employment and whether someone is enrolled in postsecondary schooling. We define employment as whether someone has positive earnings in a quarter. We measure postsecondary enrollment on the basis of state records identifying whether an individual was enrolled in courses at a state 2- or 4-year college or university.

Although our data pertain to a single state, this state is quite typical of the United States. The industrial structure is similar to that of the United States as a whole, and earnings and wages are within 10 percent of the US average. The proportion of the population that is African American is slightly below the national average. The proportion Hispanic is substantially below the US average but similar to that of most states.

Table 1 provides a tabulation of the GED scores on the first test taken and an indicator of whether the test was later retaken, for individuals tak-

[5] We draw on interviews with state agency personnel in our description of the state GED data and procedures.

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

[7] As discussed later, the regression sample used in our main analysis below is limited to individuals with test scores between 1500 and 3000, whereas samples in table 1 and in figs. 1–3 include scores outside this range.

TABLE 1
TEST PERFORMANCE AND TEST RETAKING:
FIRST-TIME TEST TAKERS, 1995–2005

Score Range[a]	Number	Distribution (%)	Retake (%)
0–990	1,009	1.0	65.7
1000–1490	897	.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
2750–3090	18,173	18.8	.4
3100–4000	5,630	5.8	.2
Total	96,883	100.0	16.1

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

ing the exam for the first time in the period of our study, 1995–2005. 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 an RD design will provide an estimate of the impact for those near the threshold, individuals whose test performance is substantially below the median. If GED certification impacts for this group are substantially different from those for other GED recipients, our measures may not be representative, although policy makers seem particularly concerned about the impact of the GED on low-skill individuals.[8]

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 (col. 3) 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. The large majority of scores that satisfy the GED passing criteria with the first test are not relevant for the RD analysis because they are far from the passing threshold. Column 3 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

[8] Previous work has argued that the labor market benefits of the GED are greater for individuals with low cognitive abilities (Murnane, Willett, and Tyler 2000; Tyler, Murnane, and Willett 2003), although the evidence is not entirely consistent.

threshold (those with total scores of 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, that is, 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 2 years of first taking the test. Second, when we look at the effect of GED certification on quarterly earnings, employment, and postsecondary enrollment, 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 twelfth 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 2-year period—that determines GED certification. Consequently, the final test score is an obvious candidate for a conventional RD 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.[9]

The discontinuity in the density for the final test score is extraordinary. The log discontinuity is close to 1.05, implying that the density to the right of 2250 is nearly three times that immediately to the left, a difference that is statistically significant at better than the 0.1 percent level (i.e., $p < .001$). 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.

[9] These methods correspond to those recommended by McCrary (2008).

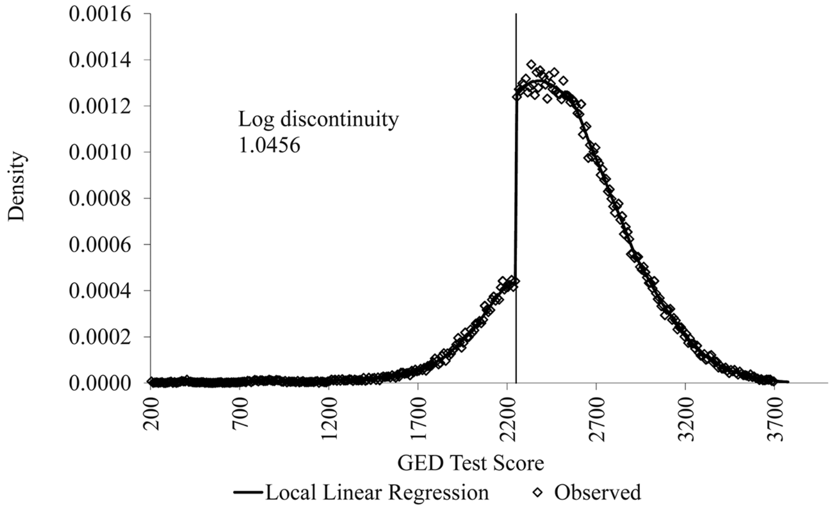

FIG. 1.—Distribution of last test score: 1995–2005

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 from those who do not, perhaps reflecting different noncognitive traits, causing those with scores just above the threshold to differ systematically from those below. Such differences might well be associated with measured personal characteristics. 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:

$$X = \alpha_x + \alpha_{xr} D_r + \sum_{j=1}^{4} \{\beta_{xlj}[D_l(\text{score} - 2250)]^j$$

$$+\beta_{xrj}[D_r(\text{score} - 2250)]^j\} + \vartheta.$$

The term D_r (D_l) is a dummy variable indicating whether that score equals or exceeds (is below) the passing threshold, and score is the final score on the GED test. The variables β_{xlj} and β_{xrj} are estimated coefficients that capture the relationship between the GED score and the dependent variable, and the coefficient α_{xr} identifies the extent of any discontinuity, providing an estimate of the mean difference on the dependent variable between those just above the threshold and those just below. The model is fitted on the sample of test scores between 1500 and 3000.

Table 2 (col. 1) provides estimates for this parameter, where the variable X is one of the following: gender (male), race (nonwhite), age,

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	−.032	−.073	−.020
	(.019)*	(.024)**	(.015)
Observations	85,402	71,854	86,345
Nonwhite:			
Coefficient	.033	.030	−.008
	(.015)**	(.018)*	(.012)
Observations	85,402	71,854	86,345
Age:			
Coefficient	−.571	−.975	−.068
	(.336)*	(.420)**	(.276)
Observations	85,402	71,854	86,345
Retake test:			
Coefficient	.047		−.484
	(.013)**		(.009)**
Observations	85,402		86,345
Prior earnings:			
Coefficient	−77.58	−1.14	−20.48
	(91.95)	(105.23)	(72.59)
Observations	85,402	71,854	86,345

Note.—Standard errors are in parentheses. The sample includes test scores from 1500 to 3000. Prior earnings are measured in the quarter before the GED attempt. The sample in col. 1 is the set of individuals taking the test for the last time between 1995 and 2005; col. 2 is the subset of individuals in col. 1 who take the test only once. The sample in col. 3 is the set of individuals taking the test for the first time between 1995 and 2005.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

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 for the final test score. The proportion of nonwhites is approximately 3 percentage points higher above the threshold than below, a difference that is easily statistically significant at the 5 percent level. The proportion of males is 3.2 percentage points lower above the threshold, and those just above the threshold are also slightly younger, differences significant at the 10 percent level. Finally, we see that those above the threshold are more likely to have retaken the test. This reflects the fact that many individuals exceed the threshold by virtue of taking the test more than once.[10]

[10] As a robustness check for our methods, we tested for discontinuities at the median for all scores below the threshold and at the median for all scores above. Of the 24 coefficient estimates to identify discontinuities for the four demographic measures, only one was significant at the 5 percent level, and one was significant at the 10 percent level.

It is clear that the final test score fails basic specification tests for a running variable (see Imbens and Lemieux 2008; McCrary 2008). The central assumption of the RD model will be violated insofar as observed differences in demographic variables imply a discontinuity in the outcome of interest (Hahn, Todd, and van der Klaauw 2001). Although it would be possible to control for these measures, a more important source of bias is potential differences in unmeasured factors, which could well be strongly associated with the outcomes of interest. For example, those who are most willing to retake the test could be individuals with noncognitive traits that make them most likely to succeed in the labor market. If this were the case, those with scores just below the threshold—disproportionately individuals who chose not to retake the test—might well have lower earnings than those just above the threshold because of such differences. Standard RD methods would mistakenly identify this difference as due to GED certification.

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. Lofstrom and Tyler (2008) limit their sample in this way as a robustness check for their RD estimation approach. 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 marked discontinuity is present just as in the final test score. This similarity indicates that the discontinuity identified in the final score in large part reflects depletion of scores just below the threshold, because individuals with these scores

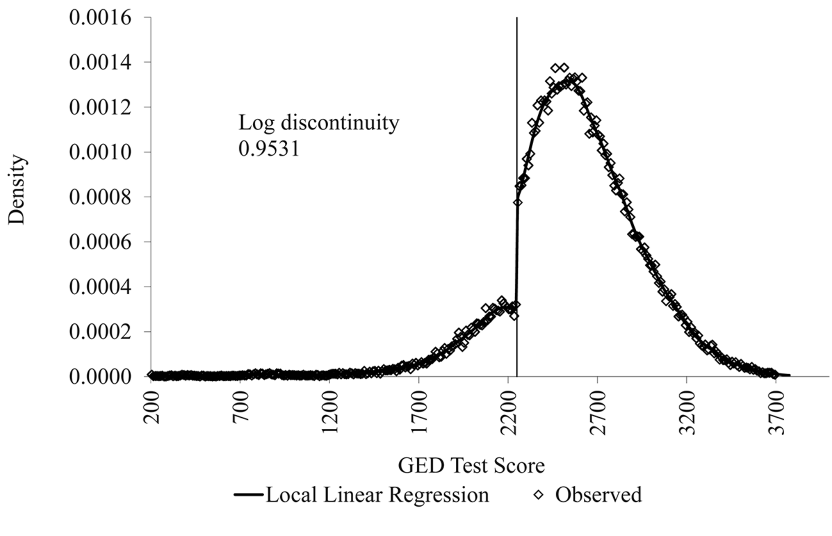

FIG. 2.—Distribution of first test score for single test takers: 1995–2005

are very likely to retake the test. Table 2 (col. 2) indicates that discontinuities in the demographic variables exist with this score and that patterns are similar to those for the final score.

An alternative is to use the first test score as the continuous variable underlying GED certification. As noted above, GED certification is not predicted perfectly by the first score. Some individuals who score below the threshold retake the test and pass it, and some who score above the threshold do not meet the subtest requirement. However, 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.[11] The assumptions underlying the FRD design imply that the first test score will display continuous relationships with all preexisting individual characteristics. Table 2 (col. 3) shows that there is no discontinuity in the demographic measures for individuals around this threshold. Figure 3 presents the distribution of the first test score, using the same method to identify discontinuities as for the densities above. The figure shows 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 an FRD design.

IV. Applying FRD Methods

Because individuals at or above the test threshold are appreciably more likely to receive GED certification than those below, these data are appropriate for an FRD design. The equation predicting GED certification is[12]

$$\begin{aligned}\text{GED} = \alpha_g + \alpha_{gr}D_r + \sum_{j=1}^{p} \beta_{glj}[D_l(T-2250)]^j \\ + \sum_{j=1}^{p} \beta_{grj}[D_r(T-2250)]^j + X\delta_g + \varepsilon,\end{aligned} \tag{1}$$

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 vector with the following set of covariates: earnings in the four quarters prior to first GED attempt, age, age squared, race, year of first GED test, quarter of

[11] We refer to this as an FRD because there are individuals whose score on the first test is below the threshold but eventually obtain a GED as well as individuals whose score is above the threshold but fail to obtain a GED. This latter group consists of individuals who fail to obtain a passing score on all five subsections of the test. Our discussion in Sec. IV includes more details on both of these groups.

[12] The formal model presented here follows closely from that in Imbens and Lemieux (2008).

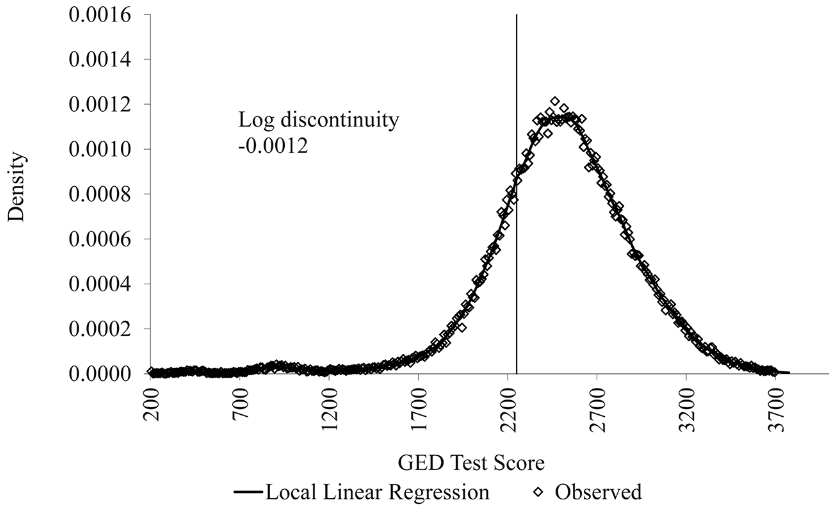

FIG. 3.—Distribution of first test score: 1995–2005

the year (winter, spring, summer, or fall), and dummies for the year in which the first test was taken. An observation is an individual who takes a first GED test. The subscript g identifies coefficients predicting GED certification, so β_{glj} and β_{grj} identify the relationship of the first GED test score with certification, below and above the 2250 threshold, respectively. The estimated parameter α_{gr} indicates the discontinuity at the threshold.

Fitting the same structure predicting the outcome variable, we write

$$\begin{aligned} Y = \alpha_y + \alpha_{yr} D_r + \sum_{j=1}^{p} \beta_{ylj}[D_l(T-2250)]^j \\ + \sum_{j=1}^{p} \beta_{yrj}[D_r(T-2250)]^j + X\delta_y + \mu. \end{aligned} \tag{2}$$

The estimate of program impact is based on the relative size of the RD 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 as

$$\tau = \alpha_{yr}/\alpha_{gr}. \tag{3}$$

Figure 4 provides a graph that illustrates the estimation methods underlying equations (1) and (2).[13] Here the focus is on earnings in quarter 12. The discontinuity assumed in equation (1) is clearly present in the data, confirming that those who score at or just above the threshold

[13] The figure shows the results from the specification that excludes covariates. The figure based on a regression specification including covariates has the same pattern as fig. 4.

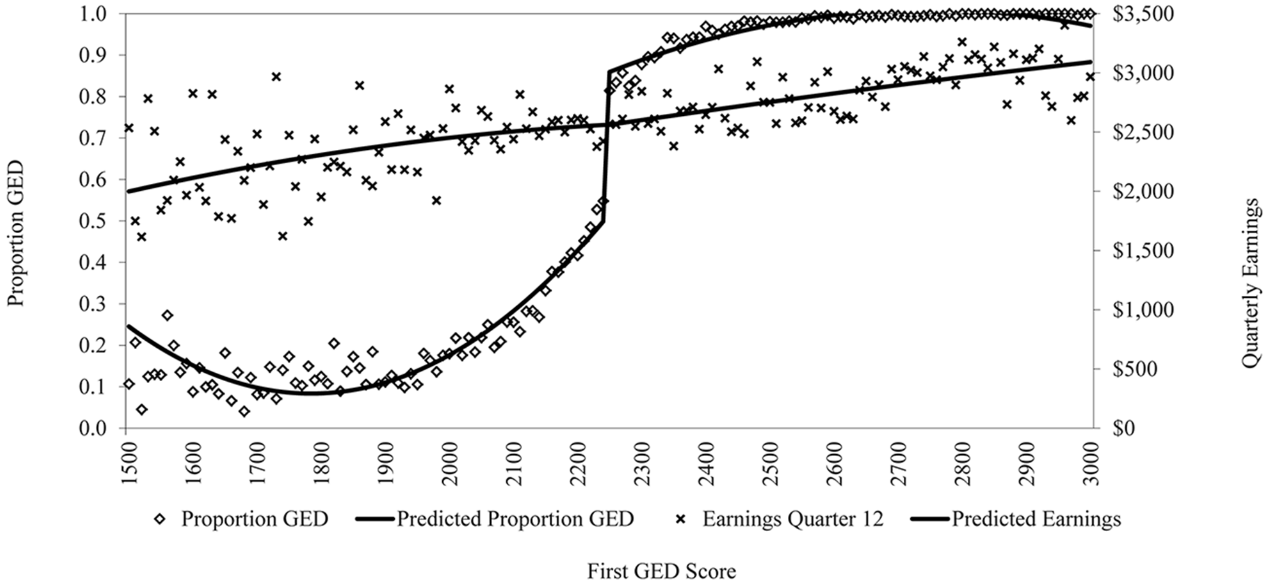

Fig. 4.—Regression discontinuity models predicting GED and quarterly earnings, men

on the overall GED score are appreciably more likely to have a GED 12 quarters after taking the test than those scoring just below. The graph for earnings does not show a discontinuity at this point, suggesting that there is little impact on quarter 12 earnings.

As Hahn et al. (2001) observe (see also Imbens and Lemieux 2008), the FRD can be formulated as a parametric instrumental variables (IV) 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:

$$\begin{aligned} Y = \ & \alpha + \tau\widehat{\mathrm{GED}} + \sum_{j=1}^{p} \beta_{lj}[D_l(T - 2250)]^j \\ & + \sum_{j=1}^{p} \beta_{rj}[D_r(T - 2250)]^j + X\delta + \vartheta, \end{aligned} \tag{4}$$

where $\widehat{\mathrm{GED}}$ is the predicted value from equation (1).[14] Because the polynomial is of the same order as in equations (1), (2), and (4), estimates of τ based on equations (1)–(3) are numerically identical to those based on equations (1) and (4). Hence, in either case, it is the estimated discontinuities in the treatment and outcome variables that determine the impact estimate, τ. Since estimating equations (1) and (4) as a parametric IV system is simpler and allows us to estimate standard errors using standard statistical software, this is how we estimate our main results from the FRD model. The reported standard errors for the second-stage equation reflect the fact that the GED measure in the equation is estimated. As a robustness check, we also use alternative methods based on linear versions of equations (1) and (2), which we describe below.

In common with other estimates based on RD methods, the validity of estimates depends on the assumption that all factors other than GED certification that are associated with the GED test score and influence outcomes are captured by our function of test score in the range of the threshold. If such factors were to change discontinuously at the threshold, this would not be captured by our smooth polynomial and would therefore induce bias in estimated GED effect estimates. We doubt that such factors are important, however, since employees have no way to provide employers with information on their scores other than GED certification itself.

The FRD provides an estimate of program impact for a subset of the population with an initial test score near the threshold, those whose GED certification is determined by whether they are above or below

[14] Estimated coefficients and error terms differ from those reported in eq. (2). Consequently, the subscript notation has been altered to reflect this.

the threshold, often referred to as “compliers.” Estimates for α_{gr} reported in table 3 (below) indicate that about a third of test takers near the threshold are compliers. Among those who satisfy the subtest requirement on the first test, compliers are those who would not retake the exam even if their score was just below the threshold and those who would retake it but would not pass the threshold.[15] Among those who do not satisfy the subtest requirement, compliers are those who would retake the test and satisfy the subtest requirement only if their overall first test score was above the threshold.

Our tabulations show that those who pass the subtest requirement but choose not to retake the test make up about 40 percent of compliers, those who pass the subtest requirement and retake the test without passing the threshold make up about 20 percent, and those who do not pass the subtest requirement but who retake the test and pass the subtest requirement if their overall score is above the threshold make up about 40 percent of compliers.[16] Note that for those with initial scores just below the threshold, table 1 indicates that 31 percent choose not to retake the test, implying that such individuals are slightly overrepresented among compliers.

Although we do not know whether there are differences in GED impact for compliers as compared to others with scores at the cutoff point, in their review of the literature, Heckman et al. (2014) find few important differences by measured characteristics aside from the results showing that women who obtain a GED are more likely to obtain employment. As noted above, some researchers have argued that lower-ability individuals have greater returns to the GED (Murnane et al. 2000; Tyler et al. 2003), but Heckman, Humphries, and Mader (2011) reject this claim. In the discussion that follows, we assume that our estimates identify the effects of the GED for all recipients.

As noted above, our basic sample includes individuals who first take the GED test in 1995–2005. We exclude those who first take the test in 2006–8 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 (many of these individuals do not take all the subtests). Substantive conclusions were not altered by this truncation, although precision of estimates was somewhat improved. This approach eliminated 8 percent of the cases below the threshold and 12 percent of the cases above the threshold. About 37 percent of those with scores

[15] In contrast, noncompliers are individuals who retake the test and pass the threshold.

[16] The first two are calculated from the sample of individuals with initial test scores in the range 2200–2240, just below the threshold; the third is calculated as a residual.

below 1500 ultimately obtain certification; virtually all those with scores over 3000 are certified. 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.

We estimate a separate regression for each of our three dependent variables in each quarter after 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. An analysis of employment is insightful when studying disadvantaged populations with low employment levels such as GED test takers. The final measure is an indicator of whether the individual enrolled in public postsecondary education in the state at any time during the quarter, based on records maintained by the state. Earnings and employment outcomes are available for 30 quarters after the initial GED attempt, whereas postsecondary enrollment is available for 16 quarters after the initial GED attempt. So we estimate 30 separate quarterly regressions for earnings and employment and 16 separate regressions for postsecondary attendance. For each equation, the order of the polynomial is two, although we also fitted higher-order polynomials and obtained similar results.

V. Results

Table 3 presents estimates based on equation (1), the first stage of the two-stage equation, applied to quarter 12.[17] In table 3, the dependent variable is a dichotomous variable for passing the GED test. Note that the first-stage estimates for the three second-stage outcomes (quarterly income, employment, and postsecondary enrollment) 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 and a 30 percentage point increase for women.

Parameter estimates for the GED impact from the basic model in equation (4), estimated separately for men and women, are presented in tables 4–6. Column 1 in each table contains the estimated impact, τ, and its standard error, as identified by the discontinuity in $\widehat{GED}$ from the first-stage equation (1). Each coefficient (and standard error) is

[17] The results from other quarters show a very similar pattern. Results from the first-stage equation differ from quarter to quarter because the dependent variable is GED certification as of the beginning of the quarter and because the sample size varies since the number of quarters of data differs by the year in which the test is taken and a small number of outliers were omitted from the earnings and employment analyses. Estimates are not affected by the latter omissions, although standard errors are reduced.

TABLE 3
First-Stage Results for Quarter 12

	Dependent Variable: GED Receipt	
	Men	Women
Discontinuity	.33984 (.00675)**	.30127 (.00663)**
Linear—left	.00190 (.00004)**	.00250 (.00005)**
Linear—right	.00061 (.00003)**	.00055 (.00003)**
Quadratic—left	.00205 (.00007)**	.00289 (.00008)**
Quadratic—right	−.00061 (.00004)**	−.00056 (.00003)**
Observations	44,378	41,967
Adjusted R^2	.5968	.6154

Note.—Standard errors are in parentheses. Controls include earnings in each of four prior quarters, nonwhite dummy, age, age squared, three dummies for the four quarters in a year, a dummy for each year the test was taken, and a constant.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

from a separate regression in which the outcome applies to the indicated quarter.

In table 4, the dependent variable is earnings for each quarter from one to 30 quarters after the initial GED test attempt. Although the estimated coefficients vary from quarter to quarter, for men, in quarters 5–9, four coefficient estimates are statistically significant at the 5 percent level (two-sided test) and one is significant at the 10 percent level. We can reject the hypothesis that the 30 coefficients are jointly zero at the 5 percent level. The largest estimated coefficient is nearly $600, or 23 percent of the mean earnings, and the other significant estimates range from $297 (12 percent of the mean earnings) to $438 (17 percent). However, no coefficient for any later quarter is statistically significant. When we estimate a model of the GED impact on total discounted earnings (with a discount rate equal to either the consumer price index [CPI] or the CPI plus 5 percent), the GED impact is 8 percent of the present value of earnings for the 30 observed quarters and is not significant at the 10 percent level (see online App. B). For women, one estimate is significant at the 10 percent level, and at $506 this amounts to 20 percent of mean earnings. However, we cannot reject the hypothesis that the 30 coefficients are jointly zero at the 10 percent level.

One obvious factor that may reduce earnings for GED recipients would be enrollment in postsecondary education. In Appendix B, we reproduced our earnings analysis limiting the sample to those not enrolled in public postsecondary education during that quarter. The results from this sample were similar 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

TABLE 4
ESTIMATED GED IMPACT ON EARNINGS

	MEN		WOMEN	
QUARTERS SINCE 1ST GED TEST	Coefficient (1)	Observations (2)	Coefficient (3)	Observations (4)
1	−25.6 (74.8)	44,378	−59.1 (62.2)	41,967
2	92.4 (106.5)	44,377	−72.5 (90.9)	41,967
3	64.6 (133.8)	44,378	45.1 (124.7)	41,965
4	69.8 (155.5)	44,378	3.1 (135.9)	41,967
5	297.5 (168.6)*	44,377	7.2 (153.4)	41,964
6	430.3 (182.5)**	44,377	120.4 (167.5)	41,967
7	429.5 (195.8)**	44,377	99.2 (183.9)	41,963
8	437.6 (207.6)**	44,376	36.9 (202.3)	41,967
9	598.7 (217.8)**	44,378	142.8 (207.4)	41,966
10	182.4 (228.1)	44,378	127.5 (217.9)	41,964
11	335.2 (234.3)	44,378	201.0 (223.5)	41,967
12	59.0 (242.8)	44,378	162.9 (236.8)	41,967
13	137.1 (249.6)	44,378	114.2 (239.5)	41,967
14	129.7 (255.0)	44,378	334.0 (245.9)	41,966
15	70.0 (266.9)	44,378	409.7 (252.3)	41,966
16	10.0 (268.7)	44,378	376.8 (256.9)	41,967
17	205.1 (272.6)	44,377	136.0 (274.5)	41,966
18	182.6 (279.3)	44,377	218.4 (268.8)	41,967
19	162.4 (283.1)	43,647	190.3 (273.1)	41,336
20	111.8 (291.5)	42,919	291.6 (282.4)	40,661
21	204.6 (303.5)	41,918	213.0 (290.5)	39,799
22	306.6 (314.0)	41,178	506.0 (307.3)*	39,092
23	295.5 (329.7)	40,483	331.2 (301.1)	38,400
24	159.8 (342.1)	39,783	378.2 (318.7)	37,726
25	493.4 (329.0)	38,966	−10.0 (319.4)	36,843
26	178.2 (332.2)	38,261	−102.4 (322.6)	36,171
27	−14.4 (338.6)	37,595	−13.5 (321.3)	35,521
28	331.2 (341.1)	36,944	136.3 (328.3)	34,860
29	135.4 (350.5)	35,983	−83.5 (329.6)	34,364
30	36.0 (358.8)	35,146	−277.7 (334.4)	33,982

NOTE.—Standard errors are in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

quarter.[18] Impact estimates in this specification were qualitatively similar to those in our base analyses.

In table 5 the dependent variable is a dichotomous variable for employment, measured as having positive earnings in the quarter. Just one of the coefficients in the table is statistically significant at the 5 percent level, and only three coefficients have absolute values greater than 0.04 (8 percent of the mean). About a third of the estimates are negative. For both men and women, we cannot reject the joint hypothesis that all

[18] Appendix B also contains a specification in which the dependent variable is the earnings level (not log), but the sample is limited to individuals with positive earnings. Again, the results are similar to those of the base analysis.

TABLE 5
Estimated GED Impact on Employment

	Men		Women	
Quarters since 1st GED Test	Coefficient (1)	Observations (2)	Coefficient (3)	Observations (4)
1	.015 (.017)	44,378	−.009 (.017)	41,967
2	−.004 (.022)	44,377	−.013 (.022)	41,967
3	.037 (.027)	44,378	−.011 (.027)	41,965
4	.010 (.029)	44,378	−.015 (.030)	41,967
5	.029 (.031)	44,377	.011 (.033)	41,964
6	.013 (.032)	44,377	.003 (.035)	41,967
7	.007 (.033)	44,377	.039 (.038)	41,963
8	.026 (.034)	44,376	.017 (.039)	41,967
9	.024 (.036)	44,378	.009 (.040)	41,966
10	.016 (.036)	44,378	.002 (.041)	41,964
11	.007 (.036)	44,378	−.014 (.042)	41,967
12	−.022 (.037)	44,378	.020 (.043)	41,967
13	−.035 (.037)	44,378	−.018 (.043)	41,967
14	.026 (.037)	44,378	.070 (.044)	41,966
15	.040 (.038)	44,378	.091 (.045)**	41,966
16	.018 (.038)	44,378	.048 (.045)	41,967
17	.014 (.038)	44,377	.017 (.045)	41,966
18	.039 (.038)	44,377	.022 (.046)	41,967
19	.022 (.039)	43,647	−.001 (.046)	41,336
20	.023 (.039)	42,919	.048 (.047)	40,661
21	−.019 (.040)	41,918	.019 (.048)	39,799
22	−.008 (.040)	41,178	.017 (.049)	39,092
23	.004 (.041)	40,483	−.014 (.049)	38,400
24	−.017 (.041)	39,783	−.006 (.050)	37,726
25	.037 (.041)	38,966	.010 (.050)	36,843
26	−.012 (.041)	38,261	.021 (.050)	36,171
27	−.003 (.041)	37,595	.016 (.050)	35,521
28	.015 (.041)	36,944	.002 (.051)	34,860
29	−.001 (.042)	35,983	−.017 (.050)	34,364
30	−.040 (.043)	35,146	.005 (.051)	33,982

Note.—Standard errors are in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

coefficients are zero. As with earnings, the results for employment are not sensitive to the inclusion of individuals attending postsecondary education during the quarter.

As mentioned above, Heckman et al. (2014) find that, in some cases, female GED recipients have higher annual earnings and employment than the full set of female dropouts. In contrast, we find similar labor market outcomes for female GED test takers whether or not they obtain GED certification. Heckman et al. show that the GED effect for women is most pronounced for the subset of women who rarely work. Since such women are unlikely to take the test, reported positive effects of the GED may actually be capturing unmeasured motivation. Our sample is limited to those who take the GED, so our comparison groups have similar levels of motivation.

Our finding that the GED affects men's earnings is at variance with most of the analyses based on comparing GED recipients and other dropouts using regression methods, which do not generally find effects. We do not emphasize these different results for several reasons. First, the positive earnings effects fade quickly, and there is no evidence that passing the GED has any long-term impact on men's earnings (the effect on the present value of earnings is not statistically significant). Second, we find no evidence of an increase in the probability of working for men in any quarter. Third, we estimate a large number of quarterly parameters, and some may be statistically significant by chance. Finally, it is useful to keep in mind that our sample differs from that of previous analyses that include both test takers and non–test takers, and effects may differ for this group.

Table 6 presents results for enrollment in the state's public postsecondary institutions. For men, GED certification is associated with increased postsecondary enrollment of 3–4 percentage points in the first three quarters after the test. We also see a positive effect of 3 percentage points in quarter 10. These are substantial effects, given that the base enrollment level is under 8 percent. In other quarters, the effect is not statistically significant. For women, the effect is even larger.[19] GED certification is associated with a statistically significant increase in the likelihood of postsecondary attendance for the first five quarters after the initial GED attempt. The effect size is 8 percentage points in the first quarter after the test, and it declines to 5 percentage points in the fifth quarter after the test, compared to a mean enrollment of 8–10 percent. In subsequent quarters, the effect continues to decline, and it is not statistically different from zero after quarter 7.

The fact that passing the GED is associated with increased postsecondary education attendance but is not associated with higher earnings or employment can be explained by the low levels of postsecondary education obtained by GED recipients. Jepsen, Mueser, and Troske (2015) show that passing the GED is associated with an increase of approximately two credit hours for men and five credit hours for women, trivially small attainment levels given that a full year of study normally entails 30 credit hours.

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 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 was nearly identical to that of the reported

[19] For both men and women, we can reject the hypothesis that the 16 GED coefficients are jointly zero at the 5 percent level.

TABLE 6
Estimated GED Impact on Postsecondary Enrollment

	Men		Women	
Quarters since 1st GED Test	Coefficient (1)	Observations (2)	Coefficient (3)	Observations (4)
1	.040 (.010)**	44,378	.081 (.012)**	41,967
2	.044 (.012)**	44,378	.067 (.015)**	41,967
3	.030 (.014)**	44,378	.042 (.018)**	41,967
4	.021 (.015)	44,378	.058 (.019)**	41,967
5	.016 (.016)	44,378	.053 (.021)**	41,967
6	.012 (.016)	44,378	.041 (.022)*	41,967
7	−.003 (.016)	44,378	.038 (.023)*	41,967
8	−.001 (.016)	44,378	.030 (.023)	41,967
9	.019 (.016)	44,378	.033 (.023)	41,967
10	.030 (.016)*	44,378	.014 (.023)	41,967
11	.019 (.015)	44,378	.020 (.023)	41,967
12	.005 (.015)	44,378	−.002 (.023)	41,967
13	−.003 (.015)	44,378	−.009 (.022)	41,967
14	.009 (.014)	44,378	−.006 (.022)	41,967
15	.006 (.014)	44,378	−.021 (.022)	41,967
16	.005 (.014)	44,378	−.004 (.022)	41,967

Note.—Standard errors are in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

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, as well as in a sample omitting 2001 and 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 RD analysis. As a robustness test, we obtained parameter estimates and standard errors 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 large as 750 points (76 data points). In no case were inferences based on these analyses seriously at variance with those presented above; nor were the impact estimates more precise.[20]

[20] This software calculates effect estimates as the ratio of estimated discontinuities, corresponding to eqq. (1)–(3), but with no covariates. As noted in the derivation of eqq. (1)–

To address the volatility of quarterly labor market outcomes for this low-skilled population, we also estimated models in which we considered aggregate measures of earnings and employment based on 1–2 years of data. Estimates of standard errors (relative to the mean) were reduced by up to 44 percent by this approach, but the substantive conclusions were not affected. For women, in no case were estimated effects on earnings statistically significant, and effects of earnings for men were not significant after year 3. We also estimated models in which employment was temporally aggregated and found that, as in the case of individual quarters, no impact estimates were statistically significant.

Our FRD model differs substantially from the RD models previously estimated for the GED. Previous work estimates a sharp regression discontinuity (SRD) based on the score obtained from a composite that includes the last test attempt of the GED. Results in table 2 show that the retaking behavior causes a discontinuity in the characteristics at the threshold for GED certification, suggesting that there may be important preexisting differences at this threshold that would bias such SRD estimates. In order to examine this possibility, we estimated SRD models of the GED using our sample of GED test takers.[21] For men's earnings, the SRD model found positive effects of the GED in later quarters, suggesting that passing the GED resulted in persistently higher earnings, whereas the FRD model showed no significant earnings gains after quarter 9. For women, the SRD estimates of earnings effects increase dramatically over the period of the study, with positive effects exceeding $400 in later quarters, most of them easily statistically significant at the 5 percent level; in contrast, estimates based on our FRD included only a single marginally significant estimate.

Differences between model estimates of GED impacts on employment were also dramatic. For men, the SRD showed negative, although usually insignificant, effects in the first six quarters and few effects in later quarters, whereas the FRD model showed no significant impacts, with estimates close to zero. For women, the SRD model estimated much larger positive employment impacts than the FRD model, with most statistically significant at the 5 percent level in later quarters. The difference in the

(4), these estimates are identical to those obtained using an IV system based on a linear functional form, but standard error estimates differ. We found that when we fitted the comparable IV model, standard errors were usually within a few percent of those obtained here, and in no case did differences influence substantive conclusions. We also estimated the comparable IV model with alternative bandwidths controlling for the same variables used in our main analyses, yielding substantively identical results. These alternative specification results are provided in App. B.

[21] Our SRD methods were designed to correspond to those of Tyler (2004). Following his procedure, we omitted cases in which the subtest requirement was not met to ensure that the test score threshold corresponded to GED certification. Results were essentially the same when we included these cases and used the final test score as a forcing variable in an FRD. Results are presented in App. B.

patterns of estimates based on the SRD and FRD for women is consistent with a positive selection story for women who take the test multiple times, similar to the selection suggested in Heckman et al. (2014).

The SRD and FRD models produced different patterns of results for postsecondary enrollment as well. For both men and women, the FRD model estimates declined over time and eventually became statistically insignificant and close to zero six to 12 quarters after the initial GED attempt, whereas the SRD results declined only slightly over time, producing statistically significant enrollment effects in every quarter for men and women.

VI. Conclusion

In this paper we have demonstrated how one can apply a valid regression discontinuity approach to a situation in which a treatment is based on a test score and individuals can influence the test score by retaking the test. We use this technique to estimate the effect of the GED test, because test takers can take the test multiple times in a 2-year period and therefore can affect their score in the immediate neighborhood of the passing threshold. We find that the effect of GED certification is small and generally not statistically significant in the long run for both men and women, but the GED is associated with a positive earnings increment in quarters 6–9 for men. We find a positive association between passing the GED and postsecondary enrollment of up to 4 percentage points for men and 8 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, and given that many postsecondary institutions require the GED or other certification, it suggests that the GED meets a perceived need for these individuals. However, these effects decline over time, becoming insignificant after the first year for men and after the second year for women. Our other work suggests that such enrollment contributes little to educational achievement.

Our results are robust to implementing the FRD technique as a local linear model or to the exclusion of demographic variables and 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.

In common with other analyses based on an RD methodology, our results formally apply only to those near the test threshold whose certification is influenced by their score relative to the threshold. It is possible that the returns to the GED for this group are lower than those of other GED recipients, but other research does not suggest that such heterogeneity is very likely. Given our use of administrative data on test takers, our results apply only to high school dropouts who are motivated enough to take the GED and may not generalize to high school dropouts who do not take the GED. Also, our analysis does not capture any direct effects of student studying on labor market outcomes. If those who study for the GED—whether or not they pass it—obtain valuable skills that improve their labor market opportunities, such benefits will not be captured in our analysis. Given that the typical GED test taker spends fewer than 40 hours studying for the test, such benefits are likely to be minimal for most test takers; but for the small group who put in substantial time studying, our approach could omit benefits they receive.

For high school dropouts near the GED passing thresholds, our findings do not support the view that GED certification is of use in helping them escape their disadvantaged labor market status. 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 to complete high school (Heckman et al. 2012). 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 body of evidence showing that GED recipients' labor market options are essentially equivalent to those of similar high school dropouts.

Appendix A

TABLE A1
Descriptive Statistics: Demographics

	Men (%)	Women (%)
Year 1st GED test:		
1995–2000	60.6	63.2
2001	13.2	13.2
2002–5	20.8	19.0
Nonwhite	21.6	19.9
GED certification	80.4	81.6
Observations	44,378	41,967

TABLE A2
DESCRIPTIVE STATISTICS: OUTCOMES

	Men					Women				
	Earnings					Earnings				
Quarters since 1st GED Test	Mean	Standard Deviation	Employment (%)	Education (%)	Observations	Mean	Standard Deviation	Employment (%)	Education (%)	Observations
1	$2,048	$2,862	62.1%	7.7%	44,378	$1,725	$2,237	63.1%	12.0%	41,967
2	$2,190	$2,980	62.6%	7.2%	44,377	$1,864	$2,345	64.1%	11.7%	41,967
3	$2,255	$3,043	62.1%	6.7%	44,378	$1,921	$2,639	64.2%	10.9%	41,965
4	$2,333	$3,131	61.7%	6.4%	44,378	$1,996	$2,484	64.6%	10.4%	41,967
5	$2,404	$3,173	61.7%	6.0%	44,377	$2,058	$2,537	64.4%	9.7%	41,964
6	$2,470	$3,251	61.1%	5.5%	44,377	$2,109	$2,550	64.3%	9.2%	41,967
7	$2,507	$3,352	60.7%	5.2%	44,377	$2,155	$2,611	64.0%	8.9%	41,963
8	$2,565	$3,374	60.5%	5.0%	44,376	$2,227	$2,733	64.0%	8.4%	41,967
9	$2,607	$3,407	60.0%	4.8%	44,378	$2,261	$2,714	63.9%	8.0%	41,966
10	$2,657	$3,515	59.4%	4.5%	44,378	$2,300	$2,790	63.5%	7.7%	41,964
11	$2,661	$3,574	58.7%	4.2%	44,378	$2,312	$2,799	62.7%	7.4%	41,967
12	$2,723	$3,625	58.3%	4.1%	44,378	$2,359	$2,913	62.5%	7.0%	41,967
13	$2,766	$3,682	57.5%	3.8%	44,378	$2,382	$2,895	62.2%	6.6%	41,967
14	$2,781	$3,728	56.9%	3.6%	44,378	$2,398	$2,899	61.9%	6.4%	41,966
15	$2,795	$3,906	56.4%	3.5%	44,378	$2,413	$2,948	61.3%	6.1%	41,966
16	$2,835	$3,830	56.1%	3.3%	44,378	$2,434	$2,974	60.8%	5.9%	41,967

17	$2,858	$3,850	55.6%	44,377	$2,457	$3,152	60.5%	41,966
18	$2,888	$3,905	55.5%	44,377	$2,483	$3,056	60.2%	41,967
19	$2,904	$3,936	55.2%	43,647	$2,473	$3,054	59.6%	41,336
20	$2,933	$3,995	54.6%	42,919	$2,477	$3,081	59.3%	40,661
21	$2,971	$4,064	54.5%	41,918	$2,500	$3,111	58.8%	39,799
22	$3,016	$4,138	54.5%	41,178	$2,527	$3,226	58.4%	39,092
23	$3,018	$4,316	53.9%	40,483	$2,517	$3,172	58.0%	38,400
24	$3,049	$4,390	53.6%	39,783	$2,555	$3,302	57.9%	37,726
25	$3,078	$4,223	53.4%	38,966	$2,566	$3,288	57.5%	36,843
26	$3,104	$4,264	53.3%	38,261	$2,598	$3,306	57.5%	36,171
27	$3,090	$4,428	52.8%	37,595	$2,576	$3,279	57.0%	35,521
28	$3,135	$4,341	52.7%	36,944	$2,604	$3,336	56.7%	34,860
29	$3,178	$4,374	52.7%	35,983	$2,604	$3,359	56.4%	34,364
30	$3,211	$4,426	52.5%	35,146	$2,633	$3,392	56.2%	33,982

NOTE.—Reported sample sizes are for the earnings and employment analyses. Samples sizes differ across quarters because follow-up data are available for a shorter period of time for more recent test takers and because a small number of earnings outliers were omitted. The sample size for the analyses of educational enrollment is slightly larger because it does not omit these outliers (see table 6).

References

Cameron, Steven V., and James J. Heckman. 1993. "The Nonequivalence of High School Equivalents." *J. Labor Econ.* 11, no. 1, pt. 1 (January): 1–47.

Cellini, Stephanie R., Fernando Ferreira, and Jess Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Q.J.E.* 125 (1): 215–61.

Fuji, Daisuke, Guido Imbens, and Karthik Kalyanaraman. 2009. "Notes for Matlab and Stata Regression Discontinuity Software." Manuscript, Dept. Econ., Harvard Univ. http://www.economics.harvard.edu/faculty/imbens/files/rd_software_09aug4.pdf.

GED Testing Service. 2013. "The 2002 Series GED Test." http://www.gedtestingservice.com/educators/current-testing-program.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201–9.

Heckman, James J., John Eric Humphries, and Tim Kautz. 2014. "The Economic and Social Benefits of GED Certification." In *The Myth of Achievement Tests*, edited by James J. Heckman, John Eric Humphries, and Tim Kautz. Chicago: Univ. Chicago Press.

Heckman, James J., John Eric Humphries, Paul A. LaFontaine, and Pedro L. Rodriguez. 2012. "Taking the Easy Way Out: How the GED Testing Program Induces Students to Drop Out." *J. Labor Econ.* 30 (3): 495–520.

Heckman, James J., John Eric Humphries, and Nicholas S. Mader. 2011. "The GED." 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., and Tim Kautz. 2014. "Achievement Tests and the Role of Character in American Life." In *The Myth of Achievement Tests*, edited by James J. Heckman, John Eric Humphries, and Tim Kautz. Chicago: Univ. Chicago Press.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *J. Econometrics* 142 (2): 615–35.

Jepsen, Christopher, Peter Mueser, and Kenneth Troske. 2015. "Second Chance for High School Dropouts? A Regression Discontinuity Analysis of Postsecondary Educational Returns to the General Educational Development Certification." Working Paper no. WP2015/08, Geary Center, Univ. Coll. Dublin.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *J. Econ. Literature* 48 (2): 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." *Res. Labor Econ.* 28:305–52.

Martorell, Paco, and Isaac McFarlin. 2011. "Help or Hindrance? The Effects of College Remediation on Academic and Labor-Market Outcomes." *Rev. Econ. and Statis.* 93 (2): 436–54.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *J. Econometrics* 142 (2): 698–714.

Murnane, Richard J., John B. Willett, and John H. Tyler. 2000. "Who Benefits from Obtaining a GED? Evidence from High School and Beyond." *Rev. Econ. and Statis.* 82 (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 the 23rd Annual Student Financial Aid Research Network Conference, Providence, RI, June 22–24.

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." Manuscript, Brown Univ.

Tyler, John H. 2004. "Does G.E.D Improve Earnings? Estimates from a Sample of Both Successful and Unsuccessful G.E.D Candidates." *Indus. and Labor Relations Rev.* 57 (4): 579–98.

Tyler, John H., Richard J. Murnane, and John B. Willett. 2000a. "Do the Cognitive Skills of School Dropouts Matter in the Labor Market?" *J. Human Resources* 35 (4): 748–54.

———. 2000b. "Estimating the Labor Market Signaling Value of the GED." *Q.J.E.* 115 (2): 431–68.

———. 2003. "Who Benefits from a GED? Evidence for Females from High School and Beyond." *Econ. Educ. Rev.* 22 (3): 237–47.