Staff Paper P18-5 December 2018

Staff Paper Series

Incentive effects of SNAP work requirements

Joseph A. Ritter

Department of APPLIED ECONOMICS

College of Food, Agricultural and Natural Resource Sciences

UNIVERSITY OF MINNESOTA

Staff Paper P18-5 **December 2018**

Incentive effects of SNAP work requirements

Joseph A. Ritter

Incentive effects of SNAP work requirements

Joseph A. Ritter*

December 17, 2018

*Department of Applied Economics University of Minnesota, St. Paul, MN 55108. jritter@umn.edu.

Acknowledgements: This research was supported by Minnesota Experiment Station Project #MIN-14-091. The author thanks Nicola Bianchi and Elizabeth Davis for helpful feedback on earlier versions.

Abstract

During the Great Recession, work requirements for various safety net programs were relaxed, and it has been argued that these contributed to high unemployment rates and long unemployment durations. One work requirement in the SNAP program applies to "able-bodied adults without dependents," and is lifted when participants reach age 50. Using a regression discontinuity approach that removes bias from age rounding, this article finds no evidence the requirement affects the probability of compliant employment when the requirement is in place.

There is perennial debate about the desirability and effectiveness of work requirements that are, or could be, attached to various welfare programs. During the Great Recession, work requirements for various safety net programs were relaxed, and it has been argued that these contributed to high unemployment rates and long unemployment durations (Mulligan, 2012). Among these programs was the Supplemental Nutrition Assistance Program (SNAP).

This study indirectly assesses this claim with respect to special SNAP work requirement for "able-bodied adults without dependents" (ABAWDs), mainly to provide evidence about the extent to which relaxing work requirements has the hypothesized unintended consequences. In addition, because work requirements add administrative burden and costs to SNAP, as well as administrative burden for recipients, it would valuable to know whether they achieve the goal of increasing labor force activity among recipients.

Since passage of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996, ABAWDs have been limited to three months of SNAP benefits unless they work at least 80 hours per month.[1] This work requirement is relaxed at age 50, providing an opportunity to evaluate its impact using a regression discontinuity (RD) design in places and times where the requirement was in force for most ABAWDs (e.g., not during the Great Recession).

Previous research has uncovered labor supply effects of the Food Stamps Program as a whole, mainly at the extensive margin. Hoynes and Schanzenbach (2012) identify labor supply impacts of the rollout of the Food Stamp program in the 1960s and 1970s using a difference-in-differences design. Consistent with standard models of labor supply, they found reduced employment (implied by the kinked budget constraint created by the program) and reduced work hours (implied by the income effect of the benefits). Fraker and Moffitt (1988), using 1980 data for female household heads, estimated that that the program reduced overall labor supply of participants by about 9 percent, but concluded that changes in specific program features had only small effects. The conclusions of these two papers are broadly similar to those for other social safety net programs such as the Earned Income Tax Credit (Eissa and Liebman, 1996; Meyer and Rosenbaum, 2001).

There was clearly a presumption among policymakers that ABAWD work requirements would offset labor supply effects of the program at the extensive margin, but this proposition has not been extensively evaluated since passage of the PRWORA. Cuffey

[1]The House version of the 2018 Farm Bill proposed to extend SNAP work requirements to older workers (between 50 and 59) and to require more hours, but the proposed extensions were dropped in negotiation with the Senate.

and Mykerezi (2016) address the incentive effects of the ABAWD work requirement in a different way than the present paper. They identify labor supply effects through difference-in-differences or triple-difference designs based on the introduction of waivers at the county level during the Great Recession using Current Population Survey (CPS) data. They find limited evidence for negative employment effects stemming from introduction of work requirement waivers.

Rather than studying introduction of waivers, the approach used here examines whether lifting the ABAWD work requirement at age 50 changes the probability of being employed at least 20 hours per week. Ideal data for this question do not exist. Therefore, I employ three different estimation samples drawn from the Current Population Survey (CPS) and the SNAP Quality Control data. Each has different strengths and weaknesses. Across all samples and all specifications, the RD analysis finds no evidence that the work requirement affects employment around age 50.

The next section describes a simple static labor supply model that illustrates the predicted effects of the work requirements. Subsequent sections discuss the data, empirical strategy, and results.

1 Theoretical labor supply implications of the work requirement

The theoretical labor supply implications of the work requirement are illustrated in Figure 1, which abstracts from benefits possibly received through other programs. In the absence of SNAP benefits an individual faces budget constraint AGEF. Introducing SNAP benefits provides income support AC for an individual who does not work. The slope of the budget constraint is smaller until benefits are completely phased out at E; the budget constraint is CEF. In a population with diverse preferences, we expect heaping near point C.

When the work requirement is introduced, SNAP benefits are not received unless the individual is working AB hours; an individual subject to the work requirement who is working less than AB hours remains on AG. An individual who is working a sufficient number of hours is on DEF, so the budget constraint is AGDE. When the work requirement is imposed, the heaping is at point D. Because of the upper age limit on ABAWD status, the prediction is that any individual who chooses D when aged 49 will prefer C after turning 50.

2 Data

The ideal sample for this study would comprise individuals whose employment could be directly affected by the ABAWD work requirements. Specifically, they would (1) choose to be SNAP participants if they could be (no never-takers), (2) meet SNAP eligibility requirements other than (possibly) employment status, and (3) meet the criteria for ABAWD status except for the possibility of being over 50 years old. For brevity I will subsequently refer to people who meet these criteria as "ABAWD-like." No data source allows us to identify ABAWD-like individuals accurately. Two strategies are feasible: (1) use a non-selective sample that includes everybody who *does* meet the criteria, but also many who do not meet any of them, or (2) use a sample of SNAP participants (who automatically satisfy criterion (1) and try to exclude as many people as possible who are not ABAWDs.

2.1 Data sources for estimation

I use both strategies in this paper. I draw a non-selective sample from the Current Population Survey (CPS). I also use samples of SNAP participants from the SNAP Quality Control (QC) data. These are described in the section 2.3. Either approach creates biases, which are discussed after the samples are described.[2]

For a RD analysis, the overall sample must also be large enough to contain a large number individuals near age 50. This rules out other possible data sources, such as the Survey of Income and Program Participation. The American Community Survey (ACS) would provide samples approximately as large as the CPS, but the reference period for the ACS question about work hours refers to the entire 12 months preceding the survey, which might include jobs or schedules that are not current. The CPS question explicitly refers to current jobs, which is more appropriate for the current purpose.

2.2 Low-waiver states

An important aspect of estimating the effect of the ABAWD work rule is to recognize that it is not always in effect in a given state. At various times states have applied

[2]Earlier versions of this paper also used a sample based on the December CPS Food Security Supplements, which ask about SNAP participation during the calendar year, linked backwards to labor force status in previous months of the year. This sample produced extremely imprecise estimates. Since they proved almost completely uninformative, I omit them in the interest of simplicity.

for and received waivers of the ABAWD work rule based on the condition of the labor market in the state or in specific counties. There is no centralized administrative source for whether these waivers were in place *and being used* by the state. Additionally, states are always allowed to exempt up to 15 percent of ABAWDs from the work rule, but states vary widely in how many of these exemptions they actually use.[3]

Because of the loose connection between availability and application of waivers and exemptions, I use the SNAP Quality Control (QC) data files to identify state-year cases where neither waivers nor exemptions were widely used. The QC files assemble annual state-by-state random samples of participating households. I categorize a state as low-waiver in a given year by first calculating the fraction of individuals aged 18 to 49, who are not living with their own children less than 18 years old, and who are coded as being in a waived area or covered under the 15 percent exemption. If the fraction is less than 5 percent, the state is considered a low-waiver state in that year and used in the analysis. Depending on the sample period, 25–29 states meet this criterion in at least one year.

The choice to use a low cutoff value rather than zero is reinforced by the fact that there are very few state-years where no ABAWD is listed as being covered by a waiver, even when other information indicates there are *no* waivers in place. The assumption implicit in using the cutoff is that the ABAWD status variables are sufficiently accurate to distinguish states in which waivers are rarely used from states in which they are common.

2.3 Details of estimation samples

The overall goal in designing the estimation samples is to get as close as possible to a sample of ABAWD-like individuals. Those under 50 would be ABAWDS whether or not they were on SNAP, and those not on SNAP would choose to enroll in the absence of the work requirement. Those over 49 would be ABAWDs in every respect except age: they would not have children under 18 or be disabled.

The non-selective estimation sample is based on the basic monthly CPS between 2000 and 2016 from the IPUMS-CPS project (Flood, King, Ruggles, and Warren, 2015). This data set is large, but there is no way to identify SNAP recipients, so I restrict the sample to individuals who did not complete high school. Lack of a high school diploma limits labor market opportunities, so the density of SNAP recipients in this sample is relatively

[3]Cuffey and Mykerezi (2016) assemble data about waivers and and use of waivers and exemptions from various sources, but they were unable to obtain information for all states.

high.[4] In the ACS for the years 2007 through 2016, between 16.5 and 29.5 percent of adults between 45 and 55 lived in SNAP households, though Meyer and Goerge (2011) estimated that 35 percent of SNAP households do not report receipt in the ACS.

The CPS no-diploma sample was restricted to U.S. citizens in a range of ages around 50, and individuals with children under age 18 were excluded. There is no way to identify other kinds of dependents or determine whether the individual is "able bodied."[5] Observations with imputed values for age or usual hours were dropped. Cases with imputed employment status were not dropped because if usual hours are not imputed, employment imputations only distinguish between being at work or not the previous week.

The second and third estimation samples use subsets of the QC data for fiscal years 2003 to 2017 and 2012 to 2017. The QC data are based on administrative review of randomly selected SNAP cases, so they accurately identify SNAP participation. The first of these QC samples identifies individuals who are not living with their own under-18 child. The second QC sample is further restricted to individuals not coded as having a disability, but this sample covers only 2012–2017 because the disability variable was introduced in 2012. The disability variable does not correspond exactly to disability status for purposes of ABAWD status, and the QC documentation warns that disability status is probably undercounted.

Figure 2 illustrates the relationship between the two QC sample definitions for 2012-2017 (the period for which they can be compared). In the age range used for estimation (45-55 years old) almost half of those who do not have a child under 18 are coded as having a disability.

2.4 Biases

Because ABAWD-like individuals cannot be fully identified in the samples, two kinds of biases arise. First, to the extent that a sample is diluted by individuals for whom

[4] This strategy of using a sample with a relatively high density of participants is similar to that used by Eissa and Liebman (1996) and Meyer and Rosenbaum (2001) to study the labor supply effects of the Earned Income Tax Credit (EITC). They chose single mothers with children as a group with a high likelihood of receiveing the EITC.

[5] Some authors exclude people living with relatives over 65, but it is not possible to identify the small fraction who are incapacitated (the term used in the SNAP regulations). CPS Annual Social and Economic Supplements (but not the monthly CPS) ask whether each adult has a disability that prevents them from working. In the 2000–2016 ASECs, only about one percent of older adults live with their 45–55 year-old children and are reported as having such a disability.

the ABAWD work requirement is irrelevant, the estimates will be attenuated. In other words, the estimates will average effects of the requirement for ABAWDs with zeroes for non-ABAWDs. Although the attenuation is not reduced in large samples, estimates from larger samples have lower variance making it more likely that an effect will be detected despite attenuation.

The CPS no-diploma sample likely has the smallest fraction of ABAWD-like individuals, but is also very large. Attenuation bias is less important in the first QC sample (no disability exclusion) because non-participants are excluded. The bias is smallest in the second QC sample which excludes some people with disabilities (but probably not all).

The second possible bias occurs in the QC samples, which include only SNAP recipients. Here estimates may be biased toward finding the hypothesized negative effect of relaxing the requirement because some people under 50, who do not or cannot meet the work requirement, lose eligibility and disappear from the data. Similarly, some of those under 50 may defer applying for benefits until after their 50$^{\text{th}}$ birthday because they do not expect to meet the work requirement.

For example, suppose that employment rates are the same before and after age 50 and consider a hypothetical sample of SNAP participants near 50 who are ABAWD-like. Leaving aside exemptions, those under age 50 will be people who are either meeting the work requirements or are in the three-month grace period. Anyone under 50 who did not find sufficient work during the three-month window will exit SNAP and not appear in the data, raising the employment rate in the sample. Since ABAWD-like people over 50 are not droppped from SNAP, they remain in the sample. Consequently, the observed employment rate is lower after age 50, making it appear that the work requirements matter for employment.

For brevity I will refer to this second type as "participation bias."[6] Participation bias is not present in the CPS no-diploma sample, but may be present in both QC samples. The overall situation is summarized in Table 1.

[6] Participation bias is a specific case of a generic concern with continuity of the the density of the forcing variable (Imbends and Lemieux, 2008). In this case the mechanism that might cause the discontinuity is clear, so the direction of bias can be assertained. Since age is verified administratively, any discontinuity of the age distribution at 50 must come from either participation rates or sampling variation.

2.5 Variables

The dependent variable in the analysis below is a dummy variable for whether the individual reports usually working at least 20 hours per week, i.e., meeting the ABAWD work requirement. Measurement of employment status and hours in the QC data may be problematic: the documentation notes caveats about employment status and hours variables ranging from moderate to strong, depending on the year. In particular, there are inconsistencies in the data suggesting that employment may be undercounted. If undercounting is a fixed percentage of employment at every age, estimates will be proportionately biased toward zero. For example, if employment rates drop from 25 percent to 20 percent at age 50, but are underreported by 10 percent, the measured drop would be from 22.5 to 18 percent, only 4.5 percentage points. However, the characteristics of the underreporting are not known. By contrast, employment is a central focus of the CPS, so it is more likely to be accurately measured.

Some estimates include state and year fixed effects, demographic controls, and education controls. The demographic controls are dummy variables for black, Asian, Native American, and Hispanic.[7] Where education controls are used, they are indicators for high school graduate, some college, bachelor's degree, master's degree, doctorate, or professional degree. Because the QC documentation recommends against using the race, ethnicity, and education variables, the corresponding controls are not used with the QC data.

3 Empirical strategy

3.1 Regression discontinuity with rounded running variable

The special work requirement for ABAWDs is lifted at age 50. If the requirement alters labor market behavior in the intended direction, there should be a fall in the compliant employment rate at age 50. To study whether the work requirement makes a difference, I employ a sharp regression discontinuity design using age as the running variable. However, reported age is rounded down in the data—an individual who is 49 years and three months is recorded as 49—which biases standard RD estimates. Therefore I use the bias correction developed by Dong (2015).

[7] I classify multi-race individuals according to the race with least labor-market advantage. For example, an individual who identifies as white and black is classified as black. Results do not change materially when only single-race respondents are included.

The adjustment works as follows. Let E_i indicate whether an individual is employed at least 20 hours per week, and suppose that a quadratic specification is assumed for $A_i = age_i - 50$. The following model is estimated:

$$E_i = \beta_0 + \beta_1 A_i + \beta_2 A_i^2 + \gamma_0 D50_i + \gamma_1 D50_i A_i + \gamma_2 D50_i A_i^2 + \varepsilon, \tag{1}$$

where $D50_i$ is an indicator of whether i is older than 50. If the distribution of birthdays within the year is uniform, the bias corrected RD estimate is

$$\hat{\tau} = \hat{\gamma}_0 - \frac{1}{2}\hat{\gamma}_1 + \frac{1}{6}\hat{\gamma}_2. \tag{2}$$

If the distribution of birthdays is not uniform, the 1/2 and 1/6 coefficients change. Taking into account the seasonality of births in 1958—approximately the midpoint of birth years for individuals aged 50 between 2000 and 2016—changes the coefficients slightly (see appendix). I do not attempt to account for seasonality of deaths, but only about 6 percent of individuals born in the relevant cohorts died by age 50 (Arias, 2014), so birth seasonality is far more important for the distribution of birthdays within the year at age 50.

3.2 Inference

Although all of the data come from low-waiver states, the strictness or leniency with of enforcement of the ABAWD work requirements may vary. Therefore most estimates use state fixed effects, and inference applies clustering at the state level. Also, the CPS uses cluster sampling.

Cameron, Gelbach and Miller (2008) recommend bootstrapping instead of standard asymptotic cluster-robust inference when the number of clusters is small or when sizes of clusters differ greatly. In the present application, the number of clusters (low-waiver states) is neither particularly small nor large (between 25 and 29 states), but the cluster sizes vary by a factor of about 50 in the CPS samples, and 15-17 in the QC samples. The variation comes partly from different size state samples in the original data and partly from the fact that states are classified as low-waiver for different years.

In testing whether lifting the work requirement has an effect, I follow Cameron, Gelbach, and Miller's recommendation to use the "wild cluster bootstrap-t imposing the null hypothesis." Confidence intervals are estimated using the procedure MacKinnon (2015) labels the "restricted bootstrap Wald interval," which is computed by inverting these hypothesis tests.[8] The procedures are described in Appendix A. The confidence intervals I report are somewhat longer than those from standard cluster-robust inference (and noticeably asymmetric in some cases).

4 Results

Figure 3 shows rates of compliant (20+ hour) employment for individuals between 45 and 55 years old in the four samples. The figure previews the formal RD results: there is no evidence in favor of the prediction of the static labor supply model. Panels (a) and (b) do not even hint that crossing the age 50 boundary affects employment rates. Employment rates are much higher in the no-diploma sample than in the QC samples (note that the scale is different), which is not surprising since many in the no-diploma sample are not eligible for SNAP benefits.

Panel (c) exhibits much more year-to-year variability due to much smaller samples. There is a fairly sharp decline between 49 and 50 for men, but the overall variation suggests this decline is likely to reflect only sampling variation; it follows a larger uptick between 48 and 49, for example.

Overall, then, there is no evidence in Figure 3 supporting the intended purpose of the ABAWD work requirement and the prediction of the static labor supply model. The remainder of the paper shows that this conclusion holds up to formal testing.

4.1 Regression discontinuity estimates

The main results reported in this section use the quadratic specification for the running variable shown in equation (1). I will show later that switching to a linear specification does not change my conclusions. The age range is 45 to 55 years, but this too does not matter for the substantive conclusions.[9]

The results mirror the casual conclusion drawn from Figure 3, so I present confidence intervals graphically for the main results to highlight the level of uncertainty about the estimated effects. Figure 4 displays point estimates and 95 percent confidence intervals

[8] Bootstrapping t statistics rather than parameter estimates also provides asymptotic refinement relative to standard cluster-robust inference, i.e., p-values are more accurate and actual coverage of the confidence intervals is closer to 95 percent than it would be using asymptotic standard errors.

[9] It would be difficult to adapt bandwidth-selection procedures for use with a rounded forcing variable.

based on the CPS sample of individuals who did not complete high school (the results are also reported in Table 2). The point estimates for men do not have the predicted sign, and every confidence interval easily covers zero.[10] The point estimates for women are very close to zero.

Clearly, the no-diploma sample provides no evidence indicating that the ABAWD work requirements have any bite near age 50, though the estimates are sufficiently imprecise that economically meaningful negative effects cannot be entirely ruled out.

The most precise estimates come from the larger QC sample comprising SNAP participants not living with children under 18. These are shown in Figure 5 and Table 3. In this case, not only are the point estimates very close to zero, but the confidence intervals suggest little uncertainty about them. These estimates may be biased toward finding the predicted negative effect (they use only SNAP participants), but also attenuated toward zero because some of those in the sample are not subject to the work requirement, e.g., are disabled.

Figure 6 and Table 4 show effects estimated from the QC sample of non-disabled individuals not living with a child under 18. This sample definition comes closest to the ideal data, but unfortunately the samples are much smaller (in addition to excluding disabled individuals, only six years of data are available). These estimates may be similarly biased toward finding the predicted negative effect and attenuated toward zero, but there is less attenuation because of removing more people not subject to the work requirement.

The top panel of Table 4 shows the only evidence in this paper suggesting support for the hypothesized effect of the work requirements: the estimates for men are both negative with p-values close to 0.1. However, this result is undermined by the falsification tests in Section 5.2: a statistically significant *positive* effect is found at age 49. The estimated effects for women are small and positive, opposite the hypothesized effect but not statistically significant.

4.2 Alternate specifications

Figures 7 through 9 vary the age range used in the estimation (i.e., the bandwidth) down to ± 2 two years and also consider a linear specification for age (corresponding tables

[10] It could be argued that clustering at the state level is too conservative, but neither this nor any subsequent conclusion changes if clustering adjustments are not used.

are in Appendix C).[11] Gelman and Imbens (2018) argue that higher order polynomials should not be used in regression discontinuity studies. Every confidence interval in the charts covers zero. The conclusion that the evidence does not support the hypothesis that the ABAWD work requirements matter is not sensitive to these variations.

5 Threats to validity

RD strategies are vulnerable to coincident discontinuities in the data around the threshold value. One such is the participation bias discussed in section 2.4. This section addresses other possibilities.

5.1 Discontinuity of covariate distributions at age 50

A general concern with the RD design is that regression samples might change in some way at age 50, masking the effect of loosening the work requirements. One way this could be indicated is discontinuity in the distribution of covariates at age 50. Therefore, I applied the same RD methodology to several of the covariates used in the CPS no-diploma samples. These were the indicators for black, white, and Hispanic (other racial groups are very small in these samples). There was no evidence of jumps in any of these variables. Because the QC documentation recommends against using the race, ethnicity, and education variables, I did not test these.

5.2 Falisification tests

Following standard practice, I to look for "treatment" effects at different values of the forcing variable, specifically looking for a discontinuity at every age between 47 and 53. Detailed results for these exercises can be found in Tables B-1 to B-3. Among six different falsification specifications each for men and women in three samples (36 regressions), nearly all estimates have large p-values.

Of particular note, however, is that the counterfactual age 49 cutoff produces a statistically significant *positive* false effect for men in the QC sample with disability exclusion

[11] Because age is measured discretely, the matrix of regressors is singular in a quadratic specification when the age range is narrower than ± 3 years. Omitting controls does not change the charts in any notable way.

(Table B-3), adjacent to the only case for which there is a statistically significant negative effect at age 50; these two results correspond to the jumps noted earlier in Figure 3, panel (c). Thus the only statistically significant estimates probably reflect only sampling variability.

Considering these alternate-age falsification tests further, note that in only one of the samples is the age-50 estimate either the smaller or the larger than all of the six estimates using counterfactual age cutoffs; in that one case it is zero to three digits. Thus none fo the actual estimates can be considered an outlier in the distribution of estimates.

6 Discussion and conclusions

This paper uses regression discontinuity methods to assess whether the ABAWD work requirements in the SNAP program actually raise employment of individuals near age 50, the point at which the requirements are loosened. I uncover no compelling evidence that they do, despite the fact that estimates from the QC samples could be biased toward finding the hypothesized effects.

Given the shortcomings of available data to address the question, a reasonable conclusion would be that the data are simply too noisy to let us see any effects that might happen when the work requirement is loosened at age 50. However, considering the results as a whole and the fact that the different samples have complementary strengths, the most reasonable conclusion is that the ABAWD work requirement has little effect on people near age 50, despite the prediction of labor supply theory and the apparent intent of Congress. Why might this be?

One conjecture is that there is asymmetry in labor supply responses: imposing a work requirement induces people to work, but when the requirement is lifted for people already working and meeting the requirement, they do not immediately adjust their labor supply. In other words, there are asymmetric effects on job finding and job separations.

A simpler possibility is that the strength of people's of labor force attachment is firmly established by the time they reach their late forties: people with a habit of employment look for and find jobs, while people who face barriers do not. Perhaps there are few people near the margin where the work requirements change behavior. If that is true, it remains an open question whether it is true for younger people.

The policy implications of this paper's findings are two-fold. First, the blanket waivers implemented during the Great Recession probably had little effect on employment, and served only to mitigate the hardship imposed by the recession (Bitler and Hoynes, 2016). Second, the resources devoted to enforcement of the ABAWD work requirement may be largely wasted.

These policy implications are, of course, subject to the caveat that the findings may not be valid for people much younger or older than 50, though it seems unlikely that people older than 50 would be more responsive to the requirements (the proposed extension of the requirements would have included those aged 50–59). Despite the limitations, the results may be compelling enough, however, to justify a randomized policy experiment such as those conducted for for unemployment insurance (Woodbury and Spiegelman, 1987; Corson et al., 1992; Spiegelman et al., 1992) and Social Security Disability Insurance (Weathers and Hemmeter, 2011).

References

Arias, Elizabeth. "United States Life Tables, 2010" *National Vital Statistics Reports,* v. 63 (2014).

Bitler, Marianne, and Hilary Hoynes. "The more things change, the more they stay the same? The safety net and poverty in the Great Recession." *Journal of Labor Economics,* v. 34 (2016).

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics,* v. 90 (2008).

Corson, Walter, Paul Decker, Shari Dunstan, and Stuart Kerachsky. "Pennsylvania Reemployment Bonus Demonstration Final Report," Unemployment Insurance Occasional Paper 92-1, U.S. Department of Labor (1992)

Cuffey, Joel and Elton Mykerezi, "Impact of SNAP work requirements: evidence from the CPS." Working paper (2017).

Dong, Yingying. "Regression discontinuity applications with rounding errors in the running variable." *Journal of Applied Econometrics,* v. 30 (2015).

Eissa, Nada and Jeffrey B. Liebman. "Labor supply response to the Earned Income Tax Credit," *Quarterly Journal of Economics,* v. 111 (1996).

Flood, Sarah, Miriam King, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business and Economic Statistics,* v. 36 (2018).

MacKinnon, James. "Wild cluster bootstrap confidence intervals," *L'Actualité économique,* v. 91 (2015).

Meyer, Bruce D. and Robert M. Goerge. "Errors in survey reporting and imputation and their effects on estimates of food stamp program participation," U.S. Census Bureau Discussion Paper CES 11-14 (2011).

Mulligan, Casey B., “Recent marginal labor income tax rate changes by skill and marital status,” NBER Working Paper 18426 (2012).

Public Health Service, U.S. Department of Health, Education, and Welfare. *Vital Statistics of the United States 1958, Volume 1.* United States Government Printing Office, Washington DC. 1960.

Woodbury, Stephen A. and Robert G. Spiegelman, “Bonuses to workers and employers to reduce unemployment: randomized trials in Illinois” *American Economic Review,* v. 77 (1987).

Spiegelman, Robert G., Christopher J. O’Leary, and Kenneth J. Kline, “The Washington reemployment bonus experiment final report,” Unemployment Insurance Occasional Paper 92-6, U.S. Department of Labor (1992).

Weathers, Robert R. and Jeffrey Hemmeter, “The impact of changing financial work incentives on the earnings of Social Security Disability Insurance Beneficiaries,” *Journal of Policy Analysis and Management,* v. 30 (2011).

Figure 1: Predicted labor supply response to SNAP work requirement

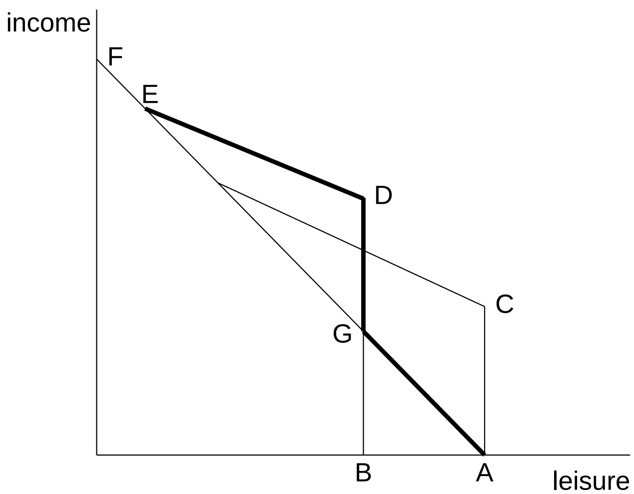
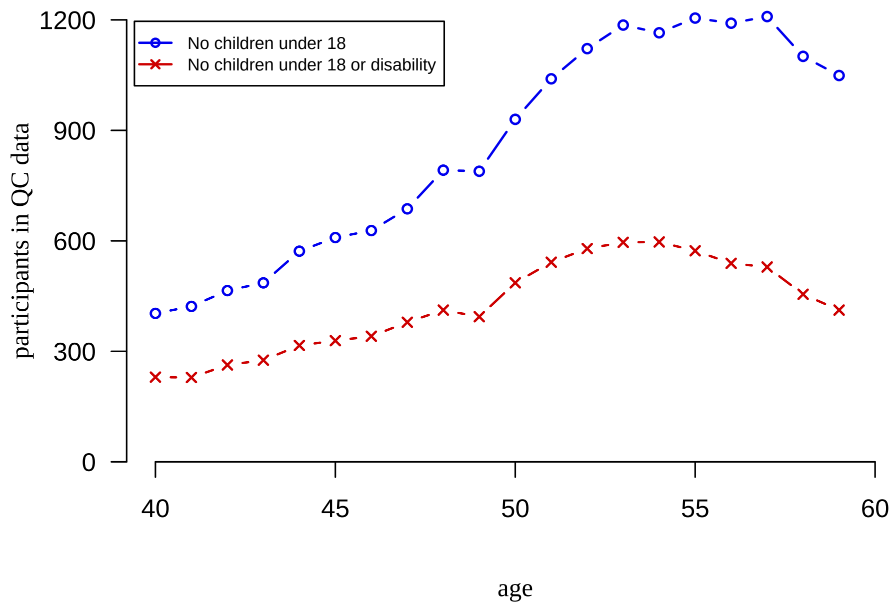


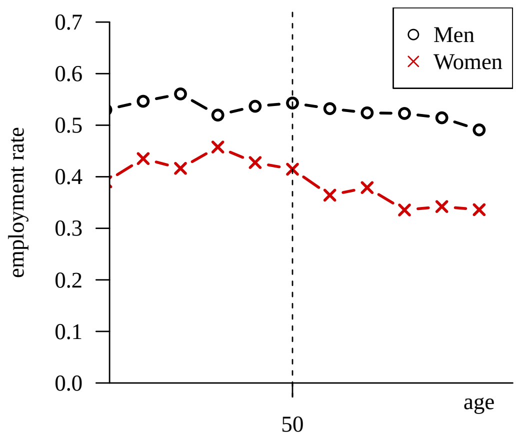
Figure 2: Age distribution of ABAWD-similar SNAP recipients in low-waiver states



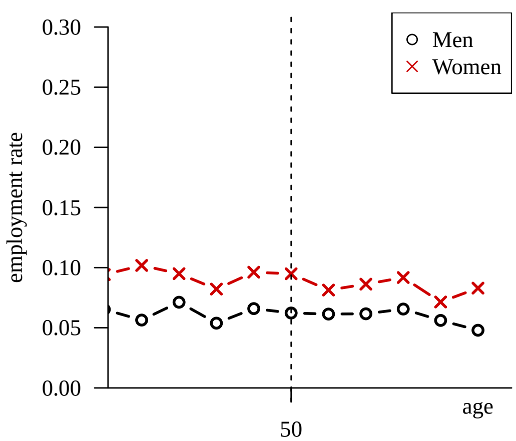
Source: SNAP Quality Control data for fiscal years 2012-2017.

Figure 3: Rates of employment at least 20 hours/week, ages 45-55

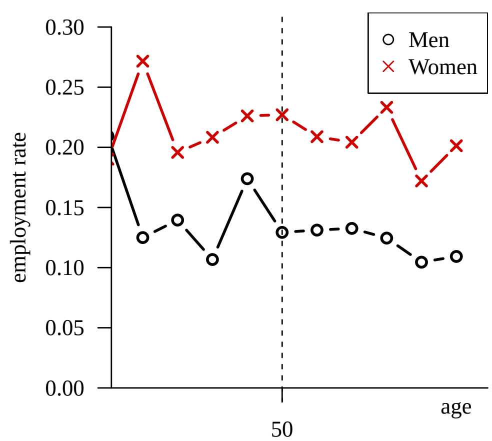
(a) CPS no-diploma sample



(b) QC sample, no disability exclusion

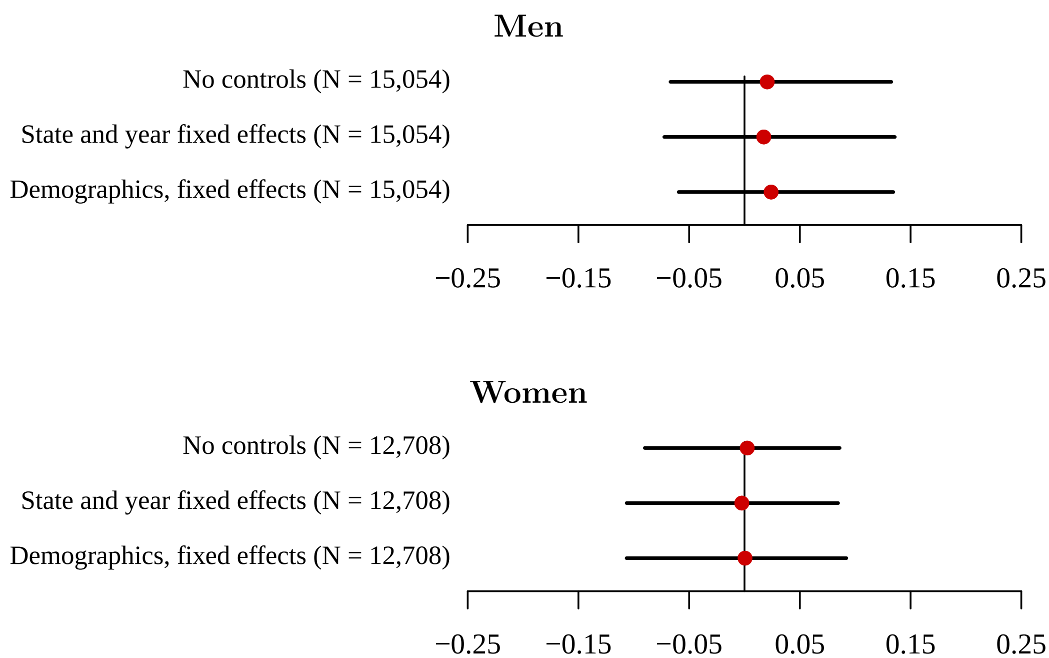


(c) QC with disability exclusion



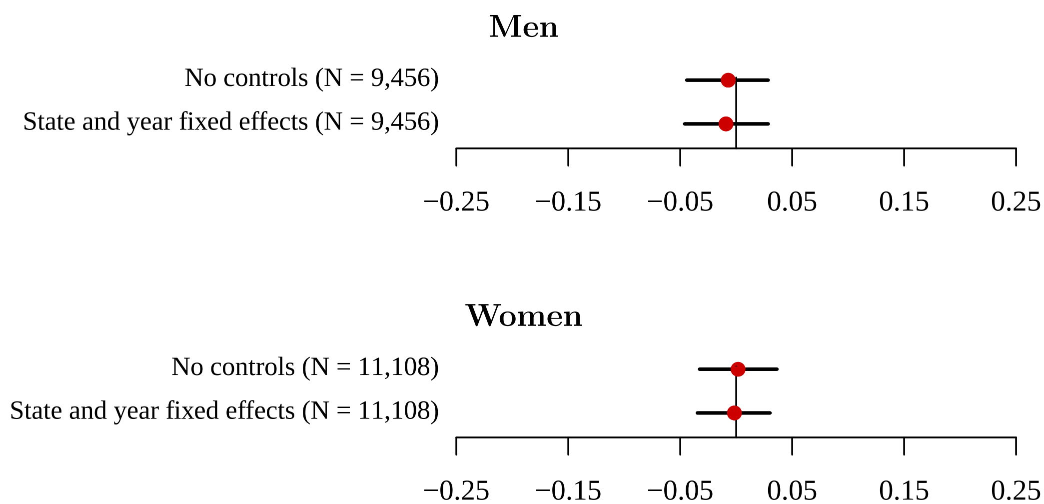
Notes: The no-diploma sample comprises individuals who did not graduate from high school. The SNAP Quality Control (QC) samples include only individuals receiving SNAP benefits. All samples are limited to individuals not living with own children under age 18 and in low-waiver states. Panel (c) also excludes individuals coded as having a disability.

Figure 4: Point estimates and 95 percent confidence intervals, no-diploma sample



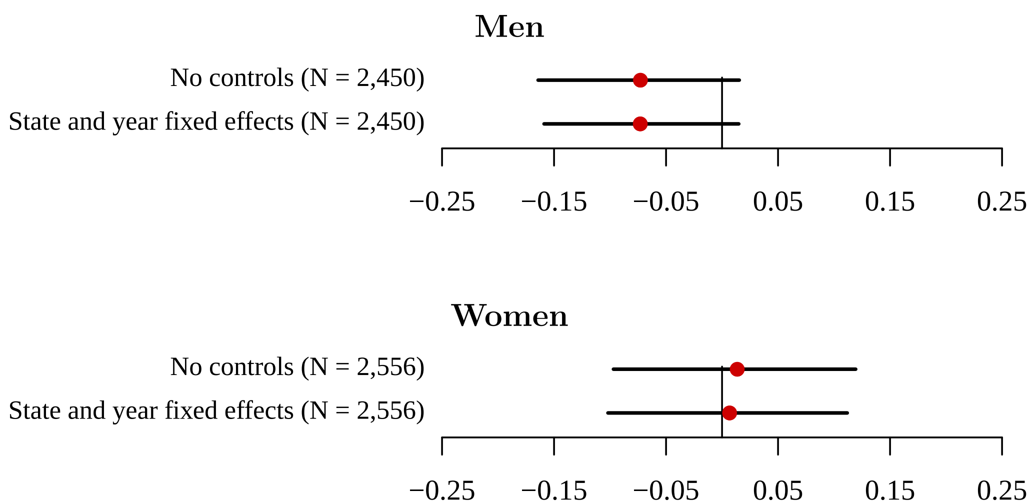
Notes: Point estimates are adjusted for rounding of age. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates. All regressions include state and year fixed effects and controls for black, Asian, Native American, and Hispanic.

Figure 5: Point estimates and 95 percent confidence intervals, QC sample without disability exclusion



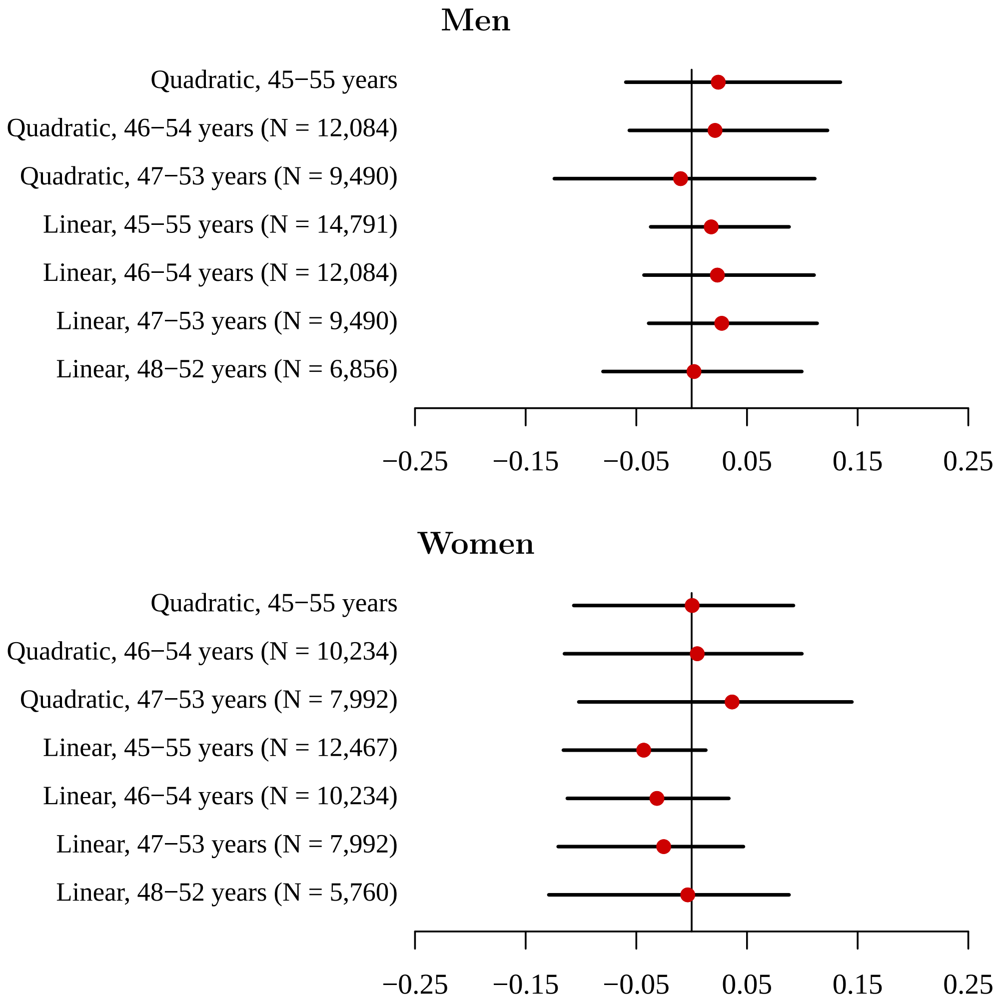
Notes: Point estimates are adjusted for rounding of age. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates.

Figure 6: Point estimates and 95 percent confidence intervals, QC sample with disability exclusion



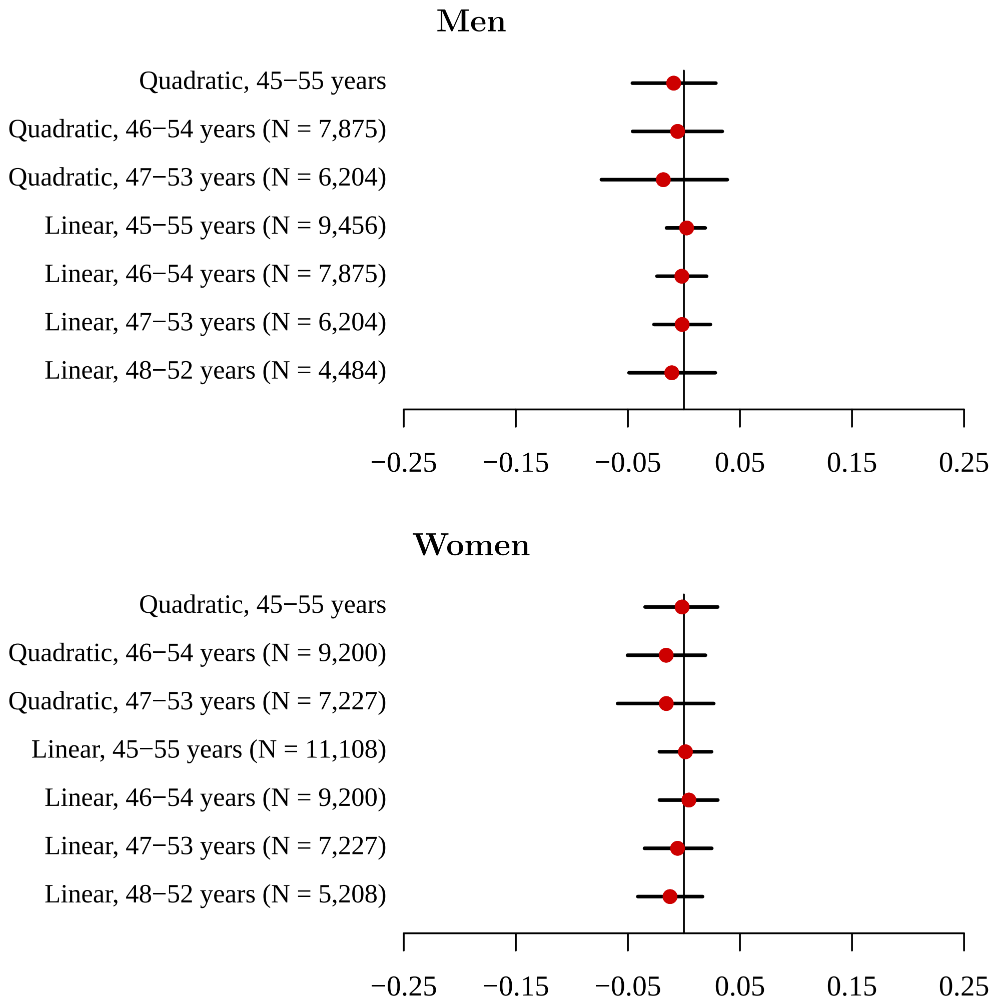
Notes: Point estimates are adjusted for rounding of age. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates.

Figure 7: Alternative estimates, no-diploma sample (95 percent confidence intervals)



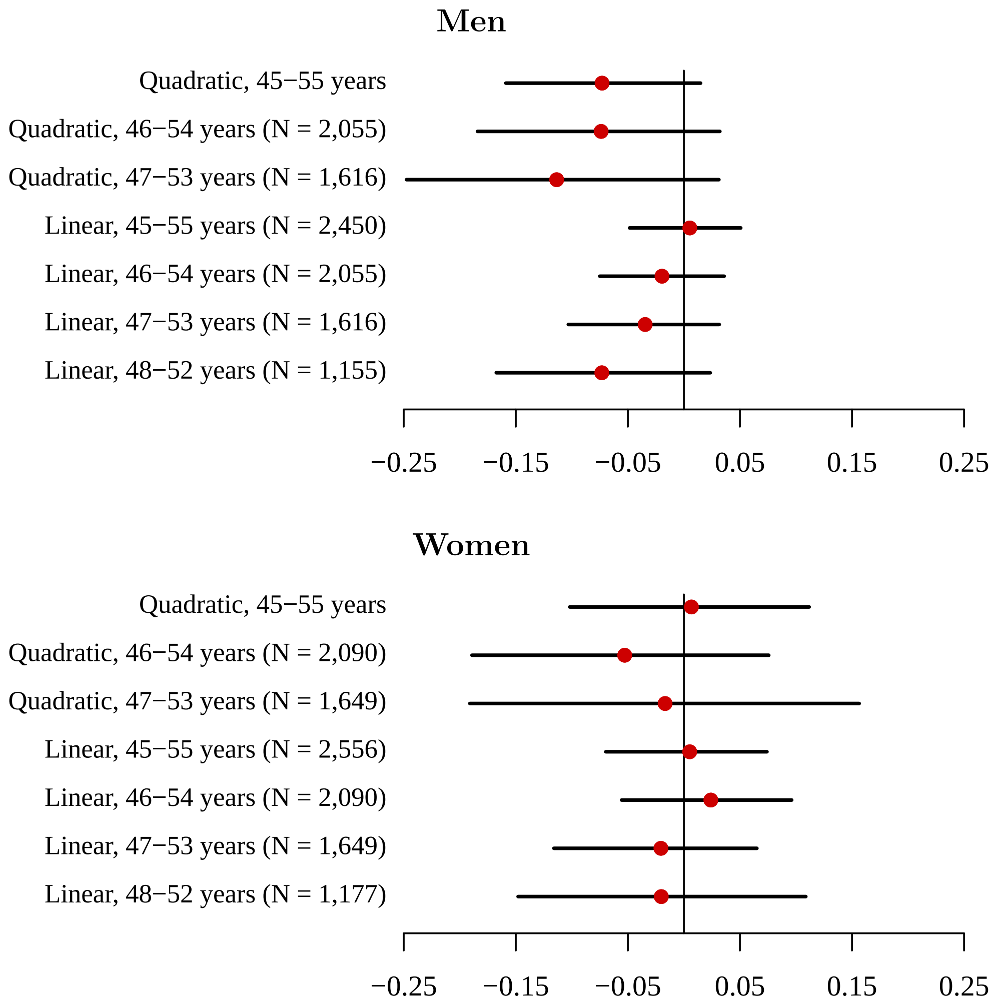
Notes: Point estimates are adjusted for rounding of age. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates. All regressions include state and year fixed effects and controls for black, Asian, Native American, and Hispanic. "Quadratic, 45-55 years" is the same as shown in Figure 4.

Figure 8: Alternative estimates, QC sample without disability exclusion
(95 percent confidence intervals)



Notes: Point estimates are adjusted for rounding of age. All regressions include state and year fixed effects. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates. "Quadratic, 45-55 years" is the same as shown in Figure 5.

Figure 9: Alternative estimates, QC sample with disability exclusion
(95 percent confidence intervals)



Notes: Point estimates are adjusted for rounding of age. All regressions include state and year fixed effects. Confidence intervals are restricted bootstrap Wald intervals, 10,000 replicates. “Quadratic, 45-55 years” is the same as shown in Figure 6.

Table 1: Summary of sample biases

Sample	Attenuation Bias	Participation bias	Sample Size (M, F)[a]
CPS, no-diploma	yes	no	15,054, 12,708
QC, no disability exclusion	reduced	yes	9,456, 11,108
QC with disability exclusion	smallest	yes	2,450, 2,556

[a]Men and women, aged 45-55.

Table 2: Employment effects, no-diploma sample

	Men		
$\hat{\tau}$	0.021	0.017	0.024
Cluster-robust SE	(0.043)	(0.044)	(0.042)
p-value†	0.663	0.710	0.564
N	15054	15054	15054
	Women		
$\hat{\tau}$	0.002	-0.002	0.000
Cluster-robust SE	(0.040)	(0.043)	(0.043)
p-value†	0.944	0.946	0.987
N	12708	12708	12708
State and year	no	yes	yes
Demographic	no	no	yes

Ages 45-55. Demographic controls are indicators for black, Asian, Native American, and Hispanic. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table 3: Employment effects, QC sample wihout disability exclusion

	Men	
$\hat{\tau}$	-0.007	-0.009
Cluster-robust SE	(0.017)	(0.018)
p-value†	0.691	0.609
N	9456	9456
	Women	
$\hat{\tau}$	0.002	-0.002
Cluster-robust SE	(0.016)	(0.015)
p-value†	0.916	0.934
N	11108	11108
State and year	no	yes

Ages 45–55. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table 4: Employment effects, QC sample with disability exclusion

	Men	
$\hat{\tau}$	−0.073	−0.073
Cluster-robust SE	(0.043)	(0.042)
p-value†	0.104	0.096
N	2450	2450
	Women	
$\hat{\tau}$	0.014	0.007
Cluster-robust SE	(0.050)	(0.049)
p-value†	0.783	0.898
N	2556	2556
State and year	no	yes

Ages 45–55. The sample excludes individuals coded as having a disability as well as individuals living with children under age 18. The disability indicator is available beginning in 2012. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Appendix A: Details about methods

Correcting for birth seasonality

Dong (2015) developed an RD estimator for use when the running variable is rounded downward to the next integer. The formula in equation (2) applies if exact age is distributed uniformly with in reported years. Births are known to be seasonal in the United States: the amplitude of seasonality in the mid-1950s was about ±8 percent. Therefore, I implement Dong's more general bias correction. For a specification that is quadratic in the running variable, Dong's formula is

$$\hat{\tau} = \hat{\gamma}_0 - \mu_1\hat{\gamma}_1 + (2\mu_1^2 - \mu_2)\hat{\gamma}_2,$$

where $\mu_1 = \mathrm{E}(e)$ and $\mu_2 = \mathrm{E}(e^2)$ and e is the rounding error.

I calibrate μ_1 and μ_2 as follows. Data on births by month are taken from the 1958 volume of *Vital Statistics of the United States* (Public Health Service, 1960). I approximate the within-year density function of births by assuming the distribution of births within each month is uniform and that the probability of birth in that month equals the proportion of 1958 births that took place during the month. The density of births within the year is the same as the density of rounding errors within the year and can be used to calculate the two expectations. The results are $\mu_1 = 0.509$ and $2\mu_1^2 - \mu_2 = 0.177$, which are used for the estimation.

Bootstrapping procedure

Adapted for this application, the wild cluster bootstrap-t imposing $H_0 : \tau = 0$ (as described by Cameron, Gelbach, and Miller, 2008) involves four steps:

1. From the main RD regression calculate $t = \hat{\tau}/\mathrm{SE}_{\mathrm{cr}}$, where $\mathrm{SE}_{\mathrm{cr}}$ is a cluster-robust standard error for $\hat{\tau}$.

2. Run the RD regression imposing $\tau = 0$ (using equation (2)) and save the residuals, $\hat{\varepsilon}^r$, and fitted values, $\hat{Y}^r$.

3. For each cluster j, form $Y_j^b = \hat{Y}_j^r + w_j\hat{\varepsilon}_j^r$ where w_j is a scalar Bernoulli random variable taking on values ±1 with probability 1/2. In other words, apply Rademacher weights at the cluster level.

4. Run the unrestricted RD regression using Y^b as dependent variable, saving $t^b = \hat{\tau}^b / \mathrm{SE}^b_{cr}$.

5. Repeat steps 3 and 4 10,000 times. The bootstrap p-value is the largest α such that $t \notin \left(t^b_{[\alpha/2]}, t^b_{[1-\alpha/2]}\right)$, where $t_{[q]}$ is the qth quantile of the t^b values.

Confidence intervals were estimated by inverting hypothesis tests using wild cluster bootstrap t-statistics imposing H_0, as described in MacKinnon (2015). Candidate endpoints were used as null hypotheses in the procedure described above. A bisection algorithm was used to find the endpoint for which the p-values were sufficiently close to 0.05. Specifically, the relevant tail of the bootstrap-t distribution was required to contain between 248 and 252 out of 10,000 replicates or the difference between the bracketeting values in the bisection algorithm was less than 0.00005.

Appendix B: Falsification tests

Table B-1: Counterfactual age thresholds, no-diploma sample

	Age cutoff (± 5 years bandwidth)						
	47	48	49	50	51	52	53
				Men			
$\hat{\tau}$	−0.083	−0.104	0.056	0.024	0.007	0.002	0.075
Cluster-robust SE	(0.188)	(0.126)	(0.063)	(0.042)	(0.041)	(0.105)	(0.111)
p-value†	0.672	0.420	0.380	0.580	0.873	0.985	0.503
N	12816	13526	14236	15054	15754	16403	16923
				Women			
$\hat{\tau}$	−0.112	−0.095	−0.118	0.000	−0.002	−0.022	−0.150
Cluster-robust SE	(0.248)	(0.130)	(0.050)	(0.043)	(0.035)	(0.086)	(0.093)
p-value†	0.661	0.473	0.042	0.992	0.951	0.801	0.123
N	10276	11084	12011	12708	13315	13918	14585

All regressions include state and year fixed effects and indicators for black, Asian, Native American,and Hispanic. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table B-2: Counterfactual age thresholds, QC sample without disability exclusion

	Age cutoff (± 5 years bandwidth)						
	47	48	49	50	51	52	53
				Men			
$\hat{\tau}$	0.013	−0.012	0.017	−0.009	0.011	0.012	0.002
Cluster-robust SE	(0.081)	(0.041)	(0.023)	(0.018)	(0.018)	(0.026)	(0.046)
p-value†	0.890	0.761	0.465	0.608	0.536	0.665	0.958
N	8505	8887	9258	9456	9651	9823	9852
				Women			
$\hat{\tau}$	0.014	0.054	0.021	−0.002	−0.002	0.001	−0.006
Cluster-robust SE	(0.070)	(0.058)	(0.031)	(0.015)	(0.017)	(0.035)	(0.042)
p-value†	0.860	0.369	0.510	0.909	0.923	0.974	0.879
N	9633	10217	10721	11108	11489	11830	12072

All regressions include state and year fixed effects. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table B-3: Counterfactual age thresholds, QC sample with disability exclusion

	Age cutoff (± 5 years bandwidth)						
	47	48	49	50	51	52	53
				Men			
$\hat{\tau}$	−0.079	0.109	0.166	−0.073	0.036	0.028	−0.001
Cluster-robust SE	(0.271)	(0.137)	(0.081)	(0.042)	(0.043)	(0.070)	(0.106)
p-value†	0.858	0.435	0.065	0.096	0.434	0.696	0.992
N	2037	2201	2357	2450	2545	2642	2667
				Women			
$\hat{\tau}$	0.124	0.218	−0.020	0.007	0.016	0.003	0.079
Cluster-robust SE	(0.245)	(0.152)	(0.084)	(0.049)	(0.054)	(0.092)	(0.150)
p-value†	0.612	0.162	0.814	0.880	0.763	0.989	0.607
N	2101	2252	2401	2556	2659	2743	2782

All regressions include state and year fixed effects. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Appendix C: Supplementary tables

Table C-1: Alternate estimates, no-diploma sample

	quadratic in age		linear in age			
50 years …	±4 yr.	±3 yr.	±5 yr.	±4 yr.	±3 yr.	±2 yr.
			Men			
$\hat{\tau}$	0.021	−0.010	0.018	0.023	0.027	0.002
Cluster-robust SE	(0.039)	(0.051)	(0.028)	(0.033)	(0.033)	(0.039)
p-value†	0.613	0.873	0.527	0.501	0.421	0.972
N	12084	9490	14791	12084	9490	6856
			Women			
$\hat{\tau}$	0.005	0.037	0.018	−0.031	−0.025	−0.004
Cluster-robust SE	(0.046)	(0.056)	(0.028)	(0.033)	(0.036)	(0.048)
p-value†	0.920	0.533	0.537	0.347	0.500	0.925
N	10234	7992	14791	10234	7992	5760

All regressions include state and year fixed effects; indicators for black, Asian, Native American,and Hispanic. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table C-2: Alternate estimates, QC sample without disability exclusion

	quadratic in age		linear in age			
50 years ...	±4 yr.	±3 yr.	±5 yr.	±4 yr.	±3 yr.	±2 yr.
			Men			
$\hat{\tau}$	−0.006	−0.018	0.002	−0.002	−0.002	−0.011
Cluster-robust SE	(0.019)	(0.026)	(0.008)	(0.011)	(0.012)	(0.018)
p-value†	0.780	0.480	0.758	0.857	0.912	0.549
N	7875	6204	9456	7875	6204	4484
			Women			
$\hat{\tau}$	−0.016	−0.016	0.001	0.004	−0.006	−0.012
Cluster-robust SE	(0.017)	(0.021)	(0.011)	(0.012)	(0.014)	(0.014)
p-value†	0.369	0.441	0.900	0.702	0.707	0.383
N	9200	7227	11108	9200	7227	5208

All regressions include state and year fixed effects. p-values are based on the distribution of bootstrapped t-statistics (10,000 replicates) with $\tau = 0$ imposed.

Table C-3: Alternate estimates, QC sample with disability exclusion

	quadratic in age		linear in age			
50 years ...	±4 yr.	±3 yr.	±5 yr.	±4 yr.	±3 yr.	±2 yr.
			Men			
$\hat{\tau}$	−0.074	−0.114	0.005	−0.020	−0.035	−0.073
Cluster-robust SE	(0.052)	(0.062)	(0.024)	(0.027)	(0.033)	(0.045)
p-value†	0.165	0.110	0.829	0.480	0.315	0.122
N	2055	1616	2450	2055	1616	1155
			Women			
$\hat{\tau}$	−0.053	−0.017	0.005	0.024	−0.020	−0.020
Cluster-robust SE	(0.062)	(0.080)	(0.033)	(0.035)	(0.043)	(0.060)
p-value†	0.402	0.820	0.875	0.518	0.647	0.726
N	2090	1649	2556	2090	1649	1177

The sample excludes individuals coded as having a disability as well as individuals living with children under age 18. The disability indicator is available starting in 2012. All regressions include state and year fixed effects. *p*-values are based on the distribution of bootstrapped *t*-statistics (10,000 replicates) with $\tau = 0$ imposed.