The Perils of Relying Solely on the March CPS: The Case of Estimating the Effect on Employment of the TennCare Public Insurance Contraction

John C. Ham
National University of Singapore, IFAU, IRP and IZA

Ken Ueda
National University of Singapore

Revised July 2017

We would like to thank, without implicating, Thomas Buchmueller, Amy Finklestein, Robert Moffitt, and Lara D. Shore-Sheppard, as well as colleagues and seminar participants at the Institute for Research on Poverty, University of Maryland, the NBER, and the National University of Singapore, for very helpful comments. Ham thanks the National Institutes of Health for financial support. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health. We both thank the University of Maryland and the National University of Singapore for financial support. Of course, we take responsibility for all errors, and emphasize that our thanks to the individuals above in no way implies that they agree with the arguments in this paper.

ABSTRACT

In a recent paper, Garthwaite, Gross and Notowidigdo (2014) report large positive labor supply effects of a major contraction in public insurance coverage in Tennessee, announced at the end of 2004 and implemented in mid-2005, using data from the March CPS. These results are important given the expansions of Medicaid coverage under the Affordable Care Act and the potential for substantial Medicaid contractions under President Trump and the Republican Congress. Their results are surprising given the previous work on the employment effects of health insurance expansions, but the authors argue that these differences in estimates are due to the fact that the Tennessee program went much higher into the income distribution than the programs studied by other researchers.

In this paper we show, under reasonable parameter restrictions, that the framework used by Garthwaite, Gross and Notowidigdo (2014) only allows for estimating the lower bound on the labor supply response to the contraction, which makes their results all the more striking. However, we show next that their large estimates are the result of focusing on the March CPS in estimation. When we use their estimation strategy on a dataset based on all the months of the CPS, or a dataset based on the American Community Survey, we find much smaller, and sometimes negative, estimates of the lower bound on the labor supply response. This result holds when we use the whole data set or allow for parameter heterogeneity by hours worked, age and education. Note that compared to the March CPS, these alternative datasets offer much larger sample sizes and are not affected by seasonal factors. We then consider a number of possible explanations for the differences in the estimates, but our results continue when we consider these modifications. We attempt to distinguish between the estimates across databases using placebo tests. While these tests reject many estimates, there is still a very wide range in the surviving estimates. Hence, we conclude that, at best, we do not have good estimates of the treatment effect of interest.

1. Introduction

There is a large body of literature on the labor supply effects of Medicaid eligibility, and it is fair to say that most of the literature argues against large effects for this zero-one dummy variable.[1] For example, Yelowitz (1995), using data from the March CPS, found large labor force participation effects of being able to obtain Medicaid coverage while working for disadvantaged mothers. However, Ham and Shore-Sheppard (2005a), also using the March CPS and the Survey of Income and Program Participation (SIPP), showed that his results for disadvantaged mothers were an artifact of constraining welfare benefits and Medicaid availability to have the same coefficient. Once this constraint was relaxed, welfare benefits, but not Medicaid eligibility, continued to affect labor force participation. Further, Meyer and Rosenbaum (2001), using data from the Current Population Survey (CPS) Outgoing Rotation Group Files and from the March CPS again for disadvantaged mothers), found an important role for welfare benefits, but not Medicaid provisions, in a static model of labor force participation. Recently Finkelstein et al. (2014) found that offering individuals Medicaid coverage in the Oregon Health Experiment had essentially no effect on participation or employment; since their result is based on a randomized trial, this evidence is perhaps the strongest to date.[2]

---

[1] See Buchmueller, Ham and Shore-Sheppard (2016) for a summary of this research

[2] Note that one study that took a more sophisticated approach allowing for heterogeneous treatment effects found Medicaid effects. Specifically, Moffitt and Wolfe (1992) consider a reduced form model of employment and Medicaid participation, where an important independent variable of interest is the imputed value of Medicaid for a given family. They allow the Medicaid effect to vary across families and find that that the value of Medicaid matters for families with high expected medical expenses.

In a recent paper, Garthwaite, Gross and Notowidigdo (2014, hereafter referred to as GGN), reported large labor supply effects for childless adults due to a major contraction in Medicaid coverage in the Tennessee program TennCare, using data from the March CPS. Specifically, in a program change announced at the end of 2004 and implemented in mid-2005, childless adults were no longer eligible for Medicaid coverage. GGN argued that their results were not necessarily inconsistent with the above studies, for example with Finklestein et al. 2014, since TennCare had much wider coverage before the contraction relative to Oregon's Medicaid program. Specifically, in Tennessee individuals could have earnings up to four times the poverty line and be eligible for Medicaid before the contraction, while eligibility in Oregon was restricted to individuals below the poverty line.

In fact, 35% of the individuals on TennCare had earnings between 100-200% of the poverty line, and 5% had earnings above 200% of the poverty line. Hence, GGN were considering a very different population than Finklestein et al. 2014, so their argument that as a result their results are likely to be different from Finklestein et al. is credible. However, it is fair to say that their results imply by far the biggest labor supply response to Medicaid seen in the literature. A specific prediction of their paper was that the Affordable Care Act would have large labor supply effects once it was in place, since a major innovation here was to cover childless adults, albeit with incomes only up to 138% of the poverty line.[3] Furthermore, GGN's estimates also predict large

[3] Again see, e.g., Buchmueller, Ham, and Shore-Sheppard (2016) for a discussion of the changes brought about by the. Affordable Care Act.

potential employment effects if President Trump and the Republican Congress repeal the large Medicaid expansion made possible by the Affordable Care Act.

In this paper, we first show that the TennCare treatment effect on labor supply in GGN's model is not point-identified, as they are estimating a reduced form employment equation. However, we also show that, under reasonable assumptions, that they estimate a lower bound for the labor supply response to the contraction. This of course makes their large estimated coefficients even more important. However, we show next that their large estimates are the result of focusing on the March CPS in estimation, and when we use their estimation strategy on a dataset based on all the months of the CPS, or a dataset based on the American Community Survey, we find much smaller, and sometimes negative, estimates of the TennCare employment effect. This result holds when we use the whole data set or allow for parameter heterogeneity by hours worked, age and education, as GGN do. In other words, the results are consistent within the various data sets. We note that compared to the March CPS, these two datasets offer much bigger sample sizes and are not affected by seasonal factors.

We then investigate possible explanations for the differences in the estimates due to i) seasonality in the March data; ii) differential sampling of urban and rural residents; iii) differences in the earlier years of the American Community Survey from the other data sets. However, none of modifications considered in this sensitivity analysis differ from our basic results.

We attempt to distinguish between the basic estimates from different datasets using placebo tests. While these tests reject many estimates, there is still a very wide range in the surviving estimates between, for example, the large triple difference (hereafter DDD) estimates from the March CPS and the much smaller, and sometimes significantly negative, DDD estimates based upon the American Community Survey.[4]

The outline of the paper is as follows. In the next section, we briefly review the TennCare contraction and the GGN approach to estimating its effect on labor supply. In Section 3 we examine the identification of the TennCare effect on labor supply in GGN's framework and show that it is only possible to estimate a lower bound for this parameter. In Section 4 we replicate their estimating equations for not only the March CPS, but also a dataset drawn from the Basic March CPS, a dataset consisting of all months of the CPS, and a dataset based on the American Community Survey. We also consider here possible expiations for the differences in results across data sets. In Section 5 we implement placebo tests with the aim of discriminating between the wide range of estimates that our replication exercise produces. Section 6 concludes the paper.

2. The TennCare Contraction and the GGN Approach to Estimating the Effect of its Contraction on Labor Supply

2.1 The TennCare Program [5]

TennCare started in 1994 with an enrollment cap of 1.775 million people and with 12 licensed managed care organizations, which included 8 HMOs and 4 PPOs. The goal was to

[4] The confidence intervals from the DDD estimates using the March CPS and using the ACS do not overlap.

[5] For informative discussions of TennCare, see GGN, Wright (2001), Wooldridge et al (1966), Chang and Steinberg (2009), and especially Bennett (2014).

provide insurance to people who were Medicaid ineligible but were either "uninsured" or "uninsurable." The eligible pool for TennCare started with people who were rejected by private insurance plans. TennCare eventually expanded to include uninsured children (without income restrictions) age 17 and older whose parents did not have access to workplace insurance, "dislocated workers" (i.e. displaced workers in the literature), and loosened income restriction levels. As a result of these and other expansions, TennCare had very generous coverage relative to other Medicaid beneficiaries. As noted above, roughly 35% of TennCare enrollees had incomes between 100-200% of the poverty line, and 5% had incomes above 200% of the poverty line.[6]

In 2001 TennCare encountered a 342-million-dollar shortfall, resulting in their largest MCO threatening to pull out of TennCare.[7] In response, TennCare required a medical review of whether enrollees were "insurable" and began a process called "reverification," which required enrollees to schedule appointments in order to determine eligibility benefits. The result was 100,000 individuals being removed from the Medicaid rolls, and roughly 20% of the TennCare enrollees being moved from the expansion program to traditional Medicaid.

In November 2004 Governor Philip Bredesen announced that as of July 1, 2005, TennCare would no longer cover adults over 19 who did not qualify for traditional Medicaid. From December

[6] These figures are taken from Wooldridge et al (1996) and refer to 1995. Unfortunately, we cannot use existing data sets to update these numbers as none of them have income and participation in public insurance in the same year.

[7] This MCO was Blue Cross Blue Shield of Tennessee, which covered almost half of all TennCare patients at the beginning of 2001. They stated that rising costs could force them to withdraw.

2004 to June 2005, there was substantial discussion of the forthcoming July 2005 contraction in the press, and it is plausible that at least some of those who would be affected starting looking for a job with health insurance before July 2005. The disenrollment process began in July 2005.

2.2 The GGN Approach to Estimating the Effect on Labor Supply of the TennCare Contraction

Using data aggregated by state from the March CPS, GGN first estimate a difference-in-difference (hereafter DD) version of the following equation

$$L_{st} = \alpha_s + \gamma_t + \beta\{I[s = Tenn] * I[t \leq 2006]\} + v_{st}, \quad s = 1,...S \text{ and } t = 1,...,T. \qquad (1)$$

In (1), L_{st} denotes the aggregate employment rate for state *s* in year *t*, $I[\bullet]$ is an indicator function equal to 1 if the condition inside the brackets is true and zero otherwise, t=2000...,2007 and s=1...,14 (Tennessee and thirteen other Southern States).[8]

GGN note that estimating (1) will not provide a consistent estimate of if the employment equation contains state specific trends μ_{st}:[9]

$$L_{st} = \alpha_s + \gamma_t + \beta\{I[s = Tenn] * I[t \leq 2006]\} + \mu_{st} + v_{st}. \qquad (2)$$

[8] The other Southern 'states' are Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, Tennessee, Texas, Virginia, South Carolina, and West Virginia.

[9] If the states share a common trend, this will be captured by the γ_t terms.

To address this possibility, they note that the TennCare contraction affected only adults without children under 17 years old, so they turn to triple difference (hereafter DDD) estimation. They implement their DDD approach by writing for childless adults

$$L_{1st} = \alpha_{1s} + \gamma_{1t} + \beta\{I[s = Tenn] * I[t \leq 2006]\} + \mu_{1st} + v_{1st}. \tag{3}$$

while writing for adults with children

$$L_{2st} = \alpha_{2s} + \gamma_{2t} + \mu_{2st} + v_{2st}. \tag{4}$$

Subtracting (3) from (4) yields

$$\begin{aligned} L_{2st} - L_{1st} &= (\alpha_{2s} - \alpha_{1s}) + (\gamma_{2t} - \gamma_{1t}) + (\mu_{2st} - \mu_{1st}) + \beta\{I[s = Tenn] * I[t \leq 2006]\} + (v_{2st} - v_{1st}) \\ &= \alpha'_s + \gamma'_t + \mu'_{st} + \beta\{I[s = Tenn] * I[t \leq 2006]\} + v'_{st}. \quad . \end{aligned} \tag{5}$$

This will lead to consistent estimation of β under the assumption that the difference between two groups in trends is constant across states, i.e. the $\mu'_{st} = \mu'_t$ and our estimating equation is[10]

$$L_{2st} - L_{1st} = \alpha'_s + \gamma'_t + \beta\{I[s = Tenn] * I[t \leq 2006]\} + v'_{st}. \quad . \tag{6}$$

Equations (1) and (6) are GGN's equations of interest. They go to some length (on p.682) to argue that (1) and (6) represent labor supply behavior, and that β is the effect of the TennCare contraction on labor supply, as opposed to labor demand. (or a combination of both):

> "*Given the details of the reform, we interpret the employment increase to be a change in labor supply rather than labor demand. We evaluate this indirectly by studying changes in average wages, since an increase in labor supply should result in a decrease in wages.... We find a statistically insignificant 2.1 percent decrease in wages with this measure. While*

[10] The μ'_t terms will be absorbed into the time dummies.

we lack the power to detect a statistically significant change in wages, the lack of a large wage increase is consistent with a change in aggregate labor supply and not the result of an unobserved labor demand shock."

3. Identification of the Effect of the TennCare Contraction on Labor Supply

In this section, we first show that estimation of β in (1) or (6), does not provide a consistent estimate the effect of the TennCare contraction on labor supply. However, we then show that their estimate of β provides consistent estimate of a lower bound for the effect of the TennCare contraction on labor supply. To show this, we first consider the case without state specific trends, and write the structural equations for demand and supply respectively as

$$L_{st}^{\text{sup}} = \theta_{1s} + \rho_{1t} + \pi_1 w_{st} + \pi_2 \{I[s = Tenn] * I[t \leq 2006]\} + e_{1st},$$
$$L_{st}^{\text{dem}} = \theta_{2s} + \rho_{2t} + \phi_1 w_{st} + \phi_2 \{I[s = Tenn] * I[t \leq 2006]\} + e_{2st}. \tag{7}$$

In (7) θ_{1s} and ρ_{1t} represent the state and year effects respectively for the labor supply equation, while θ_{2s} and ρ_{2t} represent the corresponding effects for the labor demand equation. Further, π_1 and π_2 represent the effects of the wage and the TennCare contraction on labor supply respectively, while ϕ_1 and ϕ_2. represent the analogous effects for labor demand. We expect $\pi_1 > 0$, $\pi_2 > 0$ and $\phi_1 < 0$.[11] While GGN do not take a stand on whether the TennCare contraction

[11] Since L_{st} is measured as the fraction of the labor force employed, there is no possibility of a backward bending supply curve, and π_1 must be positive.

affects labor demand, we expect, now workers will expect employers to provide health insurance, driving up the total cost of each unit of labor for a given value of w_{st}. However, our results are not qualitatively changed if $\phi_2 = 0$ and the contraction does not affect the demand curve in (7).

Next, we solve for the model's reduced form by setting $L_{st}^{\text{sup}} = L_{st}^{\text{dem}} = L_{st}$ and solving for L_{st}. This produces

$$L_{st} = \theta_s + \rho_t + \beta\{I[s = Tenn] * I[t \leq 2006]\} + e_{st}, \tag{8}$$

where

$$\beta = \left(1 - \frac{\pi_1}{\phi_1}\right)^{-1} \left(\pi_2 - \frac{\phi_1 \pi_1}{\phi_1}\right). \tag{9}$$

Rearranging (9) implies that the effect of TennCare on labor supply is given by

$$\pi_2 = \left(1 - \frac{\pi_1}{\phi_1}\right)\beta + \left(\frac{\phi_2 \pi_1}{\phi_1}\right) > \beta. \tag{10}$$

The inequality in (10) arises because $\pi_1 > 0$, $\pi_2 > 0$, $\phi_1 < 0$ an $\phi_2 < 0$, and hence GGN's estimate of β will underestimate, on average, π_2. To see that this inequality holds if TennCare does not affect labor demand, set $\phi_2 = 0$ in (7) and hence (10) to obtain

$$\pi_2 = \left(1 - \frac{\pi_1}{\phi_1}\right)\beta > \beta. \tag{11}$$

To estimate the structural labor supply equation (7), one would need to add exogenous variables X_{st} and Z_{st} to the labor supply and demand equations respectively. If Z_{st} contains at least one variable not in X_{st} then the order condition for identification is satisfied. Assuming that the rank condition is also satisfied, then (7), and hence the labor supply effect of the TennCare contraction, is identified.[12] Unfortunately, we were unsuccessful in finding a variable that plausibly affected labor demand but not labor supply, and was not a weak instrument.[13]

Hence, our analysis suggests that the labor supply effect of the TennCare contraction is only partially identified in GGN's approach in the sense that an estimate of β provides a lower bound, on average, for π_2. There are two possible responses to this result. First, since GGN's estimate of β is large, it is quite informative about the labor supply effect of the TennCare. Second, for policy we are generally interested in the reduced form effect β. We now investigate whether we find such large estimates of β from other datasets.

[12] This discussion, of course, dates back to the Cowles Foundation, see, e.g., Hood and Koopmans (1953).

[13] If we have one excluded variable, but it is a weak instrument, the rank condition will not be satisfied asymptotically.

4. **Data Sets Used for Estimation**

We estimate equations (1) and (6) for the years GGN used, 2000-2007, for several data sets; in each case we choose the sample exactly as GGN did.[14] The first dataset we use is the CPS March Supplement (MCPS). Since this is also the data GGN use, we are essentially trying to replicate their results, and find that we can do this relatively well. (They post a data appendix but not the actual dataset they used.) When we use the MCPS, we have approximately 250,000 micro observations across all the years before aggregating the data to the state level (to replicate their difference-in-difference estimates), and aggregating to the state-child/state-childless level (to replicate their triple difference estimation).

Our first *additional* dataset is drawn from the March Basic Monthly File (BMCPS), which consists of 167,000 microdata observations across all the years. The MCPS contains the respondents from the BMCPS, as well as other respondents from March, as well as a few respondents from other months.[15]

Next, to obtain our second *additional* dataset we pool all of the Basic CPS monthly surveys within a calendar year to form an 'All CPS' dataset (AllCPS). This gives us 2.06 million micro observations over 2000 to 2007, which is (not surprisingly) a substantially larger sample than either of the MCPS or BMCPS. To obtain our final *additional* data set, we draw from the 5% sample of

---

[14] As we discuss in some detail below, we restrict the sample to those who did not achieve more than a bachelor's degree, were between 21 and 64 years old (inclusive), and were not in the military.

[15] See https://cps.ipums.org/cps/basic_asec.shtml .

the 2000 decennial Census data, the 2001-2004 pilot American Community Survey, and the 2005-2007 American Community Survey to construct an ACS data set. This data set contains 3.04 million micro observations across 2000-2007, which makes it largest dataset (in terms of micro observation) used in this analysis. The main difference between the 2005-2007 ACS and the 2001-2004 ACS pilot data is that the 2005-2007 ACS has a representative sample from every county, while this is not true of the 2001-2004 data. Also, the 2001-2004 data do not allow us to distinguish between rural and urban residents, which will affect one of our sensitivity tests below.[16]

We impose the GGN sample restrictions on each dataset as follows. In each of the datasets, education, age, and occupation are available, so we can ensure that the individual is between 21 and 64, and does not have an advanced degree, and is not in the military (using occupation). We use the "age of the youngest own child in household" variable in the ACS, AllCPS, BMCPS and the MCPS. For all datasets, we use the state of residence to determine whether the respondent lived within Tennessee or in one of the other southern states.

We have already noted that below we obtain quite different results over the data sets. We note that there are potentially important differences between the databases that may explain the differences in results. First, the ACS and AllCPS, while of course the MCPS and BMCPS survey people solely during March. Thus, the latter two datasets may be affected by seasonal factors, and below we investigate whether seasonality in construction and manufacturing employment is

[16] We also explored using the Survey of Income and Program participation, but the questions in it are quite different, and hence noncomparable, to those in the data sets just described.

responsible for the differences in the results across data sets. We find that seasonality cannot explain the differences in the results across data sets.

Further, another difference between the March CPS data sets and the annual data sets arises from the fact that the TennCare contraction was implemented in July 2005. Given this, it should be considered as a comparison year for the March CPS data sets (in the absence of anticipation effects), which is how GGN treat it. 2005 is clearly a treatment year for the annual data sets, although we would expect TennCare effect to be smaller in 2005 since the treatment only covers the second half of 2005 (again in the absence of anticipation effects). We investigate this issue below by dropping the 2005 observations in all our data sets, but perhaps, surprisingly this does not qualitatively affect our conclusions.[17]

Another difference is that in 2000 and 2005-2007, the ACS is more likely to draw from metropolitan areas relative to the AllCPS, MCPS and BMCPS.[18] This clearly cannot explain the differences in the results between the March CPS data and the AllCPS data, but could explain differences with the ACS data. Below we investigate the importance of this issue by using only individuals who lived in an urban setting. We find that differences in urban/rural sampling cannot explain the differences in the results across data sets.[19]

[17] If we do not drop 2005 when considering this issue, we need to find a way for parameterizing a treatment that occurs only for half a year when we use the annual data sets.

[18] We cannot determine the rural/urban split in the 2001-2004 ACS pilot data, since they do not have urban/rural information. Hence, we drop these years in the sensitivity test discussed in this paragraph.

[19] This approach necessitates dropping the 2001-2004 ACS observations.

Finally, we would rule out differences in how the ACS and CPS ask about employment status as a potential explanation for the different results across the data sets. As Vroman (2003) and Kromer and Howard (2011) discuss, the CPS asks 4 questions: '*LAST WEEK, did you do ANY work for (pay/either pay or profit); LAST WEEK, did you do any unpaid work in the family business or farm?; LAST WEEK, did you have more than one job/job or business, including part time? Altogether, how many jobs/jobs or businesses did you have?* In contrast, the ACS for the years we used asked only one question: *LAST WEEK, did this person do ANY work for either pay or profit?* (Kromer and Howard 2011). The CPS, as a result, captured more unpaid work and therefore had a slightly higher employment rate. If the TennCare contraction caused individuals to leave unpaid work for paid work in order to obtain health insurance, this will be *not* being counted as an increase in employment in the CPS, but it will be in the ACS.This suggests that estimates based on the CPS data sets could produce smaller treatment effects than estimates based on the ACS data. We do not pursue this given that the ACS treatment effects are generally smaller than the estimated effects from the CPS. Further, this cannot explain differences between the March CPS data sets and the AllCPS. Finally, only 1000 respondents in the entire AllCPS data set say that they only have unpaid work.

5. Empirical Results

5.1 Basic Results

In Panel A of Table 1 we present the baseline difference-in-difference (hereafter DD) estimates for the treatment effect of the Medicaid contraction on employment based on equation

(1). In the first column, we have copied the results from GGN; thus, we repeat their DD estimate (standard error) for the TennCare effect) of 2.5 (1.1) percentage points (hereafter PP). In column (2), we present our estimates *from* the MCPS, and find that for all practical purposes we replicate their results, since we find a 2.2 (1.0) PP estimated effect increase.[20] In column (3) we report the results for the BMCPS, and obtain an estimated treatment effect of 2.0 (1.1) PP. We start to diverge from the GGN results when we use the AllCPS data. Now column (4) indicates that we estimate a significant, but considerably smaller, treatment effect of 1.3 (0.4) PP. (This estimate is about half of GGN's estimated treatment effect.) We diverge even further from GGN's results when we use the ACS, since column (5) reports that we estimate a significantly *negative* treatment effect of -1.1 (0.4) PP.

Panel B of Table 1 presents the triple difference (hereafter DDD) estimates for the different datasets. The first column contains GGN's estimate of 4.6 (2.0) PP. Columns (2) shows our estimated treatment effect for the MCPS as 5.5 (2.0) PP, so again we essentially replicate their DDD results. The BMCPS estimate in column (3) is 7.0 (2.4) PP. As in GGN, the estimated DDD treatment effect is at least double the DD estimated effect when we use any variant of the March CPS. For the AllCPS, the DDD estimate in column (4) is 1.3 (0.4) PP, quite similar to the DD estimate in column (4) of Panel A of 1.4 (0.8) PP respectively. However, now the AllCPS DDD estimate is between one-third and one-quarter of size of the DDD estimates from the March CPS datasets. The estimated DDD effect from the ACS in column (5) is 0.2 (0.6) PP, and is quite small and insignificant. It is well

[20] Recall that their actual data are not available online.

known that the DD and DDD estimates will differ if there are different linear trends for the comparison and treatment groups. In Table 1 the main differences are between the DD and DDD estimates for the March CPS datasets, although there is also change in the ACS coefficient too. However, testing the hypothesis that the respective coefficients are equal is not straight-forward, and there is substantial overlap in the DD and DDD confidence intervals for the MCPS and for the BMCPS.[21]

5.2 Allowing for Heterogeneous Treatment Effects

GGN also estimate treatment effects (separately) across by i) hours worked; ii) age and iii) education groups. The motivation for differences across hours worked is that if individuals are entering employment to find private insurance to replace TennCare, they are much more likely to find such insurance in full time jobs. One explanation of why treatment effects by age and education is that take-up behavior may differ across demographic groups, as found by Ham, Ozbeklik, and Shore-Sheppard (2015). If a group had low take-up of TennCare, it is unlikely to show a big response to its elimination. Alternatively, it may be easier for some demographic groups to find a job with health insurance than others. GGN find significant treatment effects for three (not mutually exclusive) groups: i) those with more hours of work per week; ii) those with more education iii) older workers. We focus our discussion on the sensitivity of the results for these groups across data sets.

[21] If the DD estimates were efficient, we could use a Hausman test here. One way to test the DD and DDD estimates is to use placebo tests. We will show below that the March CPS DD estimates fail the placebo tests while the DDD estimates are not rejected by this test.

Table 2A contains the DD results by hours worked, while Table 2B contains the DDD results by hour worked (Note that for ease of exposition we have copied the relevant results from Table 1A and Table 1B into the first row of Tables 2A and 2B respectively). The results for those working more than 20 hours per week, and those working more than 35 hours per week, generally mimic those in Table 1A.

We have placed the respective DD an DDD treatment effects by education in Appendix Table 1, and they again follow the pattern by data set In Tables 1A and 1B respectively for those with higher education. Finally, we show the DD and DDD estimates for older and younger workers respectively in Appendix Table 2. The results once again mimic those in Tables 1A and 1B by data set. Hence the estimates for the demographic groups that GGN focus on also show substantial differences across the different data sets along the same lines as in Table 1. Hence, below we concentrate on the differences in the main results in Tables 1A and 1B below.

5.3 Trying to Explain the Difference in the Results Across Data Sets

In this sub-section, we consider differences in the data sets (discussed earlier) that may help explain the differences in the data sets. For example, probably the most glaring difference between the data sets is that the March CPS data sets do not cover July 2005, the time when the contraction started, while the annual data sets do, so we reestimate Table 1 when we omit data from 2005. We placed the results in Table 3, which indicates, perhaps surprisingly, that dropping the 2005 data does not affect our qualitative results – the only change is that the BMCPS DD estimate of the treatment effect is no longer statistically significant.

Secondly, we may be able to reduce the differences between the March datasets and the annual datasets if we remove highly seasonal employment in manufacturing and construction from each dataset. We have placed the results of doing this in Table 4, and see that these new results duplicate those in Table 1, except that now the ACS DD estimate, while still negative, is no longer significant.

Another source of the differences could be differing proportions of urban and rural workers across the datasets, combined with parameter heterogeneity across these groups. For example, the ACS and MCPS have 16.5% and 22%, respectively, of their respondents explicitly stating they do not live within a metropolitan area. We investigate whether this difference is important in Table 5 by restricting the data to include only those residing within a metropolitan area; here we need to drop the 2001-2004 ACS data since we cannot distinguish between rural and urban individuals. However, the only major difference from Table 1 is that now the AllCPS DD and DDD estimates are no longer statistically significant.

Finally, we investigate whether the 2000 decennial and the 2001-2004 ACS pilot datasets are inconsistent with the 2005-2007 ACS. While we do not expect the combination of these datasets this to be an issue, we re-ran our analysis for 2005-2007 in Table 6: the only difference from Table 1 is that the ACS DD estimate is no longer statistically significant. We note that the MCPS, BMCPS, and AllCPS coefficients, for both the DD and DDD estimation, are now larger than in Table 1, with the MCPS and BMCPS estimates being considerably larger now. This difference could be attributable to differences in the employment trends prior to TennCare repeal across the treatment and comparison groups for the March CPS data. This raises the issue of model misspecification in the different datasets, and we now turn to this issue.

6. Specification/Placebo Tests for the GGN Model Across Datasets

To carry out our specification tests, for each dataset, we run the GGN model for the years 2000-2005 while allowing for a placebo treatment that is coded 1 for 2003-2005, and 0 otherwise; as is well known, if the model is properly specified for the dataset, the placebo treatment variable should not have a significant estimated coefficient. If the placebo coefficient is significantly different from 0, its size is interesting since it gives a measure of how badly the assumption underlying the DD and DDD estimates are violated. We also repeat the analysis letting the placebo variable start to take on the value one in 2002-2005, and then the value one in only 2004-2005, respectively.

Row 1 of Table 7 contains the DD estimates of the placebo effects for the case where the placebo variable equals one for 2003 and later, for the full datasets. These estimates imply that the model is misspecified for all datasets except the ACS data, and the biases for the MCPS, BMCPS and AllCPS are considerable at -0.030, -0.025, and -0.015 respectively. Row 2 of Table 6 contains the DDD estimates of the placebo effects for the full datasets. Now only the AllCPS data produces a significant placebo effect. The corresponding figures showing the trends for the treatment and comparison group relevant for DD estimation across the datasets are in Figures 3A-3C, while the corresponding figures relevant for DDD estimation across the datasets are in Figures 4A-4C. We view these as bearing out the more formal results above, with the possible exception of the DDD figures for the March CPS, which would seem to indicate different high order trends..

We repeat the analysis for the case where the placebo variable takes on the value of one in 2002 and later in Appendix Table 3; the corresponding estimates for the case where the placebo variable takes on the value of one in 2004 and later are in Appendix Table 4. The 2002 placebo effects are significant only in the MCPS and the AllCPS data when we use DD estimation, and are insignificant for all the datasets when we use DDD estimation. When we let the placebo treatment take place in 2004 and later, they are significant with DD estimation for all datasets.

However, when we use DDD estimation, the placebo effects are significant for the AllCPS data. If we consider only the estimates that never have a significant placebo effect, we are left with the DDD estimates from the MCPS (and BMCPS) data and the ACS data. Unfortunately, the DDD estimates for the MCPS and the BMCPS at 5.5 (2.0) PP and 7.0 (2.5) PP, respectively, are very different than the DDD estimate for the ACS at -0.2 (0.6). The 95% confidence intervals for the DDD estimates from the MCPS, BMCPS and ACS are [1.5, 9.5] PP, [2.0, 12.0] PP and [-1.4, 1.0] PP, and hence the confidence intervals for the DDD March CPS estimates do not even overlap with those from the ACS. Thus, at least for the full datasets, the placebo tests do not reduce our uncertainty about the overall TennCare treatment effects.

Moreover, we consider the estimates of the 2003 placebo effect for when we let the treatment effect vary by hours worked, education, and age, respectively, in Tables 8A-8B, Appendix Table 5, and Appendix Table 6 respectively. For the case where the treatment effects differ by hours worked, we focus on the DD and DDD placebo effects in Tables 7A and 7B respectively for working more than 20 hours per week and working more than 35 hours per week respectively, where we found large treatment effects for the March CPS data. For those working more than 20 hours results, the placebo tests mimic those for the full data in Table 6: for the MCPS and BMCPS, the placebo effects are significant for DD estimation but not DDD estimation; for the AllCPS, the placebo effect is significant for both DD and DDD estimation, and the ACS placebo effect is not significant for either DD or DDD estimation. Returning to Tables 2A and 2B, We are left choosing between estimates of 4.6 (2.2) PP, 5.2 (2.6) PP, 0.4 (0.7) PP from the DDD estimation and -0.3 (0.7) PP. from DD estimation. When we consider the results for working more than 35 hours, the major difference is that the MCPS and BMCPS placebo effects are significant for both DD and DDD estimation, and we are left with only the DD ACS estimate of -1.3 (0.4) PP and the DDD ACS estimate of -0.8 (0.8) PP. Hence these tests lead us away from the GGN estimates.

Appendix Table 5 contains the estimated placebo effects by education group. Here we focus on the results for those with a high school degree or more, as that group that had the larger treatment effects in Appendix Table 1. Now the placebo tests mirror those for the full data in Table 6: for the MCPS and BMCPS, the placebo effects are significant for DD estimation but not DDD estimation; for the AllCPS, the placebo effect is significant for both DD and DDD estimation, and the ACS placebo effect is not significant for either DD or DDD estimation. Thus, returning to Appendix Table 1, we are left with estimates of 3.8 (2.1) PP, 6.9 (2.5) PP, and -0.01(0.06) PP from the DDD estimation and -0.08 (0.04) PP and -0.07 (0.09) from the DD estimation.

Finally, Appendix Table 6 contains the estimated placebo effects by age group. Again, we focus on older workers, the group that had the larger treatment effects in Appendix Table 2. Now the placebo effects are significant for the MCPS, BMCPS, and AllCPS when we use DD estimation, but none of the datasets produces a significant placebo effect when we use DDD estimation. As a result, returning to Appendix Table 2, we are left with estimates of 6.8 (2.7), 5.8 (3.3), 0.9 (1.0) PP and 0.7 (0.8) PP from the DDD estimation and -0.014 (0.005) PP from the DD estimation.

7. Conclusion

In this paper, we first show that the GGN approach does not allow for point identification of the labor supply response to the TennCare contraction. However, we also show that, given reasonable assumptions, the GGN approach, in fact estimates a lower bound for the treatment effect. As a result, GGN's large estimated effects become even more striking when compared to the previous estimates of the effect of public health insurance expansions and contractions. We reestimate the model across different datasets and find that the results are not robust to moving to a dataset based on all of the CPS months, or one based on the American Community Survey. We also find some important differences between the AllCPS results and the ACS results.

We find that these differences are not the result of the March CPS being substantially affected by seasonality in manufacturing and construction. We also find that the differences between the AllCPS and ACS are not due to the ACS sample containing more urban households than the AllCPS, or to an improper merging of the 2001-2004 ACS data with the 2000 and 2005-2007 data. Finally, we find that the differences between results for the March data and the annual data sets are not a result of misclassifying 2005 as a control year when using the annual data.

We then turn to specification tests to distinguish between the different estimates. While these tests eliminate many specifications, they do not reduce the range of the estimates since the very large DDD March CPS estimates, and the small ACS DDD estimates are not rejected. Thus, readers may differ in their estimate of the TennCare contraction treatment effect. One could argue that the ACS data dominate the March CPS data by virtue of being much larger and not subject to seasonality. In this case, one would conclude that the employment effects of the TennCare contraction are close to zero, and we learn nothing about the labor supply effect (since a low bound of zero is not informative on a parameter that should be positive). Alternatively, one could put the results from the ACS data and the March CPS data on equal footing, which would imply that the TennCare contraction, despite its large size, is essentially uninformative with regard to its employment and labor supply effects. We find it difficult to come up with convincing arguments that imply that the March CPS data dominate the ACS data.

Our paper has implications well beyond an analysis of TennCare. Many scholars understandably focus in on one data set even when there are other data sets that can be used for replication, given the effort that goes into analyzing the one data set. However, our paper shows there is a risk in proceeding this way. Moreover, we do not believe that there is evidence that the instability across data sets that we find is some sort of fluke or rare case, since even in the Medicaid literature, we have seen that Cutler and Gruber's (1996) estimated take up and crowd-out rates from the March CPS are much larger than those found by Ham and Shore-Sheppard (2005b) using the Survey of Income Participation.

References

Bennett, Christina Juris (2014). TennCare, One State's Experiment with Medicaid Expansion. Nashville: Vanderbilt University Press.

Buchmueller, Thomas, John C. Ham and Lara Shore-Sheppard. (2016). "The Medicaid Program." In the Economics of Means-Tested Transfer Programs II, edited by Robert Moffitt, NBER volume, Chicago: University of Chicago Press.

Chang, Cyril and Stephanie Steinberg (2009). "TennCare Timeline: Major Events and Milestones from 1992 to 2016," Methodist LeBonheur Center for Healthcare Economics, University of Memphis, 20.

Cutler, David and Jonathan Gruber. (1996). "Does Public Insurance Crowd Out Private Insurance?" Quarterly Journal of Economics, Vol. 111, No. 2, pp. 391-430.

Finkelstein, Amy, Katherine Baicker, Jae Song, and Taubman, Sarah (2014). "The Impact of Medicaid on Labor Market Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment." *American Economic Review*, Volume 104, Issue 5 322-28.

Garthwaite, Craig, Tal Gross, and Matthew Notowidigdo, (2014). "Public Health Insurance, Labor Supply, and Employment Lock." *The Quarterly Journal of Economics*, Volume 129, Issue 2: 653–696.

Ham, John and Lara Shore-Sheppard (2005a). "Did Expanding Medicaid Affect Welfare Participation?" *Industrial and Labor Relations Review*, Volume 58, Issue 3: 3452-470.

Ham, John and Lara Shore-Sheppard (2005b). "The Effect of Medicaid Expansions for Low-Income Children on Medicaid Participation and Private Insurance Coverage: Evidence from the SIPP." *Journal of Public Economics 2005* 89: 57-83.

Hood, W. C. and T. C. Koopmans (editors) (1953) Studies in Econometric Method. New York: John Wiley and Sons.

Meyer, Bruce and Daniel Rosembaum (2001). "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *The Quarterly Journal of Economics*, Volume 116, Issue 3: 1063-1114.

Moffitt, Robert and Barbara Wolfe. (1992). "The Effect of the Medicaid Program on Welfare Participation and Labor Supply." Review of Economics and Statistics, Vol. 74, No. 4, pp. 615-626.

Wooldridge, Judith, Leighton Ku, Teresa Coughlin, Lisa Dubay, Marilyn Ellwood, Shruti Rajan, and Sheila Hoag (1996). "Implementing State Health Care Reform: What Have We Learned from the First Year? The First Annual Report of the Evaluation of Health Reform in Five States," *Mathematica Policy Research.*

Wright, Douglas (2001). "TennCare: A Closer Look," (Tennessee Comptroller of the Treasury Legislative Briefing Paper. (2001).

Yelowitz, Aaron (1995). "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *The Quarterly Journal of Economics*, Volume 110, Issue 4: 909-939.

Table 1: TennCare Effect on Employment by Database

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Point Estimate	0.025**	0.022**	0.020*	0.013***	-0.011***
Standard Error	(0.011)	(0.010)	(0.011)	(0.004)	(0.004)
N	136	136	136	136	136
Panel B: Triple Difference					
Point Estimate	0.046**	0.055***	0.070***	0.014*	-0.002
Standard Error	(0.020)	(0.020)	(0.024)	(0.008)	(0.006)
N	272	272	272	272	272
Unconditional Average	0.705	0.705	0.707	0.705	0.709
Microdata Sample		249,559	167,368	2,057,701	3,036,337

Notes: The years used are every year from 2000 to 2007 for the MCPS, ACPS, and ACS. We use all people between 21 to 64 years old (inclusive) who were not part of the active military, with at most a bachelor's degree for the MCPS, ACPS, and ACS. We use the same set of southern states used within Garthwaite et al. (Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, Tennessee, Texas, Virginia, South Carolina, and West Virginia). We compute the share of employment within each state-year combination. Column (1) shows the same set of statistics displayed within Garthwaite et al. (2014). Column (2) shows our version of their empirical specification using the MCPS. Column (3) shows the equivalent statistics using only March from the ACPS. Column (4) shows the equivalent statistics using all months from the ACPS. Column (5) shows the equivalent statistics using the ACS. The standard errors are calculated in the same fashion as Garthwaite et al.

Table 2A: TennCare Effect on Employment (Hours) by Database: Difference in Difference

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
			Working > 0 hours		
Point Estimate	0.025**	0.022**	0.020*	0.013***	-0.011***
Standard Error	(0.011)	(0.010)	(0.011)	(0.004)	(0.004)
			0<Working < 20 hours		
Point Estimate	-0.001	-0.001	-0.004	0.000	0.001
Standard Error	(0.004)	(0.004)	(0.005)	(0.001)	(0.001)
			Working □ 20 hours		
Point Estimate	0.026***	0.024**	0.024**	0.014***	-0.012***
Standard Error	(0.010)	(0.011)	(0.012)	(0.004)	(0.004)
			Working □ 20 hours, < 35 hours		
Point Estimate	0.001	0.000	0.007	0.001	0.001
Standard Error	(0.007)	(0.007)	(0.008)	(0.002)	(0.002)
			Working □ 35 hours		
Point Estimate	0.025**	0.023**	0.018	0.014***	-0.013***
Standard Error	(0.011)	(0.011)	(0.012)	(0.004)	(0.004)

Notes: We use the same years and states as Table 1. For the MCPS, ACPS, and ACS, within each state-year combination, we compute the share of employment, those who work less than 20 hours a week, more than 20 hours a week, more than 20 hours a week and below 35 hours week, and more than 35 hours. Each row represents a different difference in difference regression where the corresponding outcome was used as the dependent variable. The standard errors are calculated in the same fashion as Garthwaite et al.

Table 2B: TennCare Effect on Employment (Hours) by Database: Triple Difference

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
		Working > 0 hours			
Point Estimate	0.046**	0.055***	0.070***	0.014*	-0.002
Standard Error	(0.020)	(0.020)	(0.024)	(0.008)	(0.006)
		0<Working < 20 hours			
Point Estimate	0.002	-0.002	0.003	-0.001	0.001
Standard Error	(0.009)	(0.009)	(0.010)	(0.003)	(0.002)
		Working > 20 hours			
Point Estimate	0.044**	0.057***	0.067***	0.015*	-0.003
Standard Error	(0.020)	(0.021)	(0.025)	(0.008)	(0.006)
		Working >20 hours, < 35 hours			
Point Estimate	0.018	0.008	0.022	0.014***	-0.008
Standard Error	(0.013)	(0.014)	(0.016)	(0.005)	(0.004)
		Working > 35 hours			
Point Estimate	0.026	0.046**	0.052**	0.004	0.004
Standard Error	(0.021)	(0.022)	(0.026)	(0.009)	(0.007)

Notes: We use the same years and states as Table 1. For the MCPS, ACPS, and ACS, within each state-year combination, we compute the share of employment, those who work less than 20 hours a week, more than 20 hours a week, more than 20 hours a week and below 35 hours week, and more than 35 hours. Each row represents a different triple difference regression where the corresponding outcome was used as the dependent variable. The standard errors are calculated in the same fashion as Garthwaite et al.

Table 3: TennCare Effect on Employment by Database – Omit 2005

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Point Estimate	--	0.019*	0.016	0.012***	-0.013***
Standard Error	--	(0.011)	(0.012)	(0.004)	(0.004)
Panel B: Triple Difference					
Point Estimate	--	0.051**	0.066***	0.012	-0.004
Standard Error	--	(0.022)	(0.026)	(0.008)	(0.006)
Microdata Size		216,751	145,633	1,789,894	2,491,229

Notes: Except for data from 2005, we use the same years and states as Table 1.

Table 4: TennCare Effect on Employment by Database - Omit Construction, Manufacturing

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
		Panel A: Difference-in-Difference			
Point Estimate	--	0.026***	0.024*	0.013***	-0.005
Standard Error	--	(0.010)	(0.014)	(0.005)	(0.005)
N					
		Panel B: Triple Difference			
Point Estimate	--	0.054**	0.074***	0.019**	-0.007
Standard Error	--	(0.024)	(0.028)	(0.009)	(0.007)
Microdata Sample		207,036	139,105	1,712,681	2,468,114

Notes: We use the same years and states as Table 1 and we impose an additional restriction: We further restrict the sample of people to those who were not part of the construction or manufacturing sectors. All other calculations are performed analogously to those in Table 1.

Table 5: TennCare Effect on Employment by Database - Only Metropolitan Areas

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
			Panel A: Difference-in-Difference		
Point Estimate	--	0.021*	0.024*	0.002	-0.010**
Standard Error	--	(0.012)	(0.014)	(0.004)	(0.005)
			Panel B: Triple Difference		
Point Estimate	--	0.044*	0.061**	0.008	0.001
Standard Error	--	(0.023)	(0.029)	(0.009)	(0.008)
Microdata Sample		193,290	128,292	1,584,874	1,822,645

Notes: We use the same years and states as Table 1 but we now further restrict the sample of people to those who were part of a metropolitan area. Sine we do not have information on whether the respondents lived in an urban or rural area in the 2001-2004 ACS, we drop those years. All other calculations are performed analogously to those in Table 1.

Table 6: TennCare Effect on Employment by Database - 2005 to 2007

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
		Panel A: Difference-in-Difference			
Point Estimate	--	0.040***	0.042**	0.017***	0.002
Standard Error	--	(0.015)	(0.017)	(0.005)	(0.004)
N		51	51	51	51
		Panel B: Triple Difference			
Point Estimate	--	0.071**	0.087***	0.022**	0.006
Standard Error	--	(0.030)	(0.034)	(0.010)	(0.009)
Unconditional Average	--	0.701	0.703	0.705	0.709
N		102	102	102	102
Microdata Sample		99,381	65,594	802,771	1,680,411

Notes: We use the same states as Table 1. We use the years 2005 to 2007 instead of 2000 to 2007. All other restrictions and calculations are done in the same manner as in Table 1.

Table 7: Placebo Tests Assuming a 2003 Treatment Year

	GGN	MCPS	BMCPS	ACPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Point Estimate	--	-0.030***	-0.025**	-0.015***	-0.001
Standard Error	--	(0.011)	(0.013)	(0.004)	(0.004)
N	102	102	102	102	102
Panel B: Triple Difference					
Point Estimate	--	-0.012	-0.036	-0.022***	-0.006
Standard Error	--	(0.021)	(0.024)	(0.008)	(0.008)
N	204	204	204	204	204

Notes: We use the same states as Table 1. We use the years 2000 to 2005 instead of 2000 to 2007. All restrictions and calculations are done in the same manner as in Table 1. We treat 2003, 2004, and 2005 as the treatment years as the placebo test.

Table 8A: Placebo Tests Assuming a 2003 Treatment Year – Treatment Effect Varies with Hours Worked; Difference in Difference

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
		Working > 0 hours			
Point Estimate	--	-0.030***	-0.025**	-0.015***	-0.001
Standard Error	--	(0.011)	(0.012)	(0.004)	(0.004)
		0< Working < 20 hours			
Point Estimate	--	-0.001	0.001	0.000	0.001
Standard Error	--	(0.005)	(0.005)	(0.001)	(0.001)
		Working > 20 hours			
Point Estimate	--	-0.029***	-0.026**	-0.014***	-0.002
Standard Error	--	(0.011)	(0.013)	(0.004)	(0.004)
		>20 hours, Working < 35 hours			
Point Estimate	--	-0.013*	-0.012	-0.006**	-0.001
Standard Error	--	(0.007)	(0.008)	(0.003)	(0.003)
		Working > 35 hours			
Point Estimate	--	-0.018	-0.020	-0.008**	-0.002
Standard Error	--	(0.012)	(0.013)	(0.004)	(0.004)

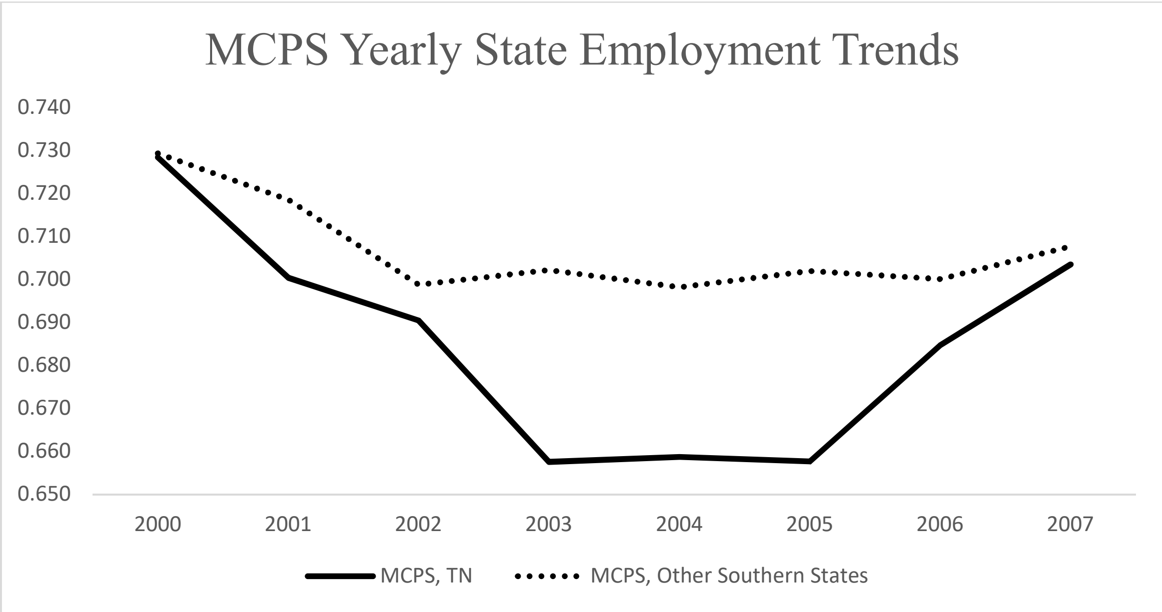
Notes: We use the same states as Table 1. The outcomes are the same as Table 2A. We use 2000 to 2005 instead of 2000 to 2007. We use 2003, 2004, and 2005 as the treatment years.

Table 8B: Placebo Tests Assuming a 2003 Treatment Year – Treatment Effect Varies with Hours Worked; Triple Difference

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
		Working >0 hours			
Point Estimate	--	-0.012	-0.036	-0.022***	-0.006
Standard Error	--	(0.021)	(0.023)	(0.007)	(0.007)
		0 <Working < 20 hours			
Point Estimate	--	0.002	0.002	-0.008***	-0.001
Standard Error	--	(0.009)	(0.010)	(0.003)	(0.002)
		Working >20 hours			
Point Estimate	--	-0.013	-0.038	-0.014*	-0.005
Standard Error	--	(0.021)	(0.025)	(0.008)	(0.008)
		Working > 20 hours, < 35 hours			
Point Estimate	--	0.020	0.023	0.005	0.002
Standard Error	--	(0.014)	(0.017)	(0.005)	(0.005)
		Working > 35 hours			
Point Estimate	--	-0.039*	-0.057**	-0.020**	-0.008
Standard Error	--	(0.022)	(0.026)	(0.008)	(0.008)

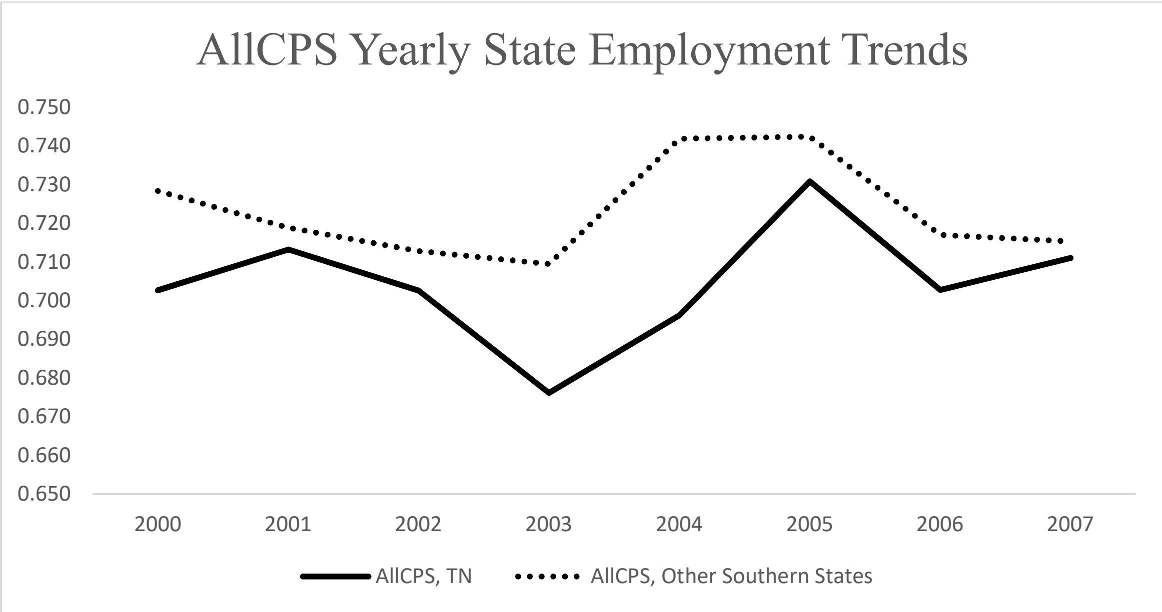
Notes: We use the same states as Table 1. The outcomes are the same as Table 2B. We use 2000 to 2005 instead of 2000 to 2007. We use 2003, 2004, and 2005 as the treatment years in the difference in difference analysis.

Figure 1A: MCPS Yearly State Employment Trends



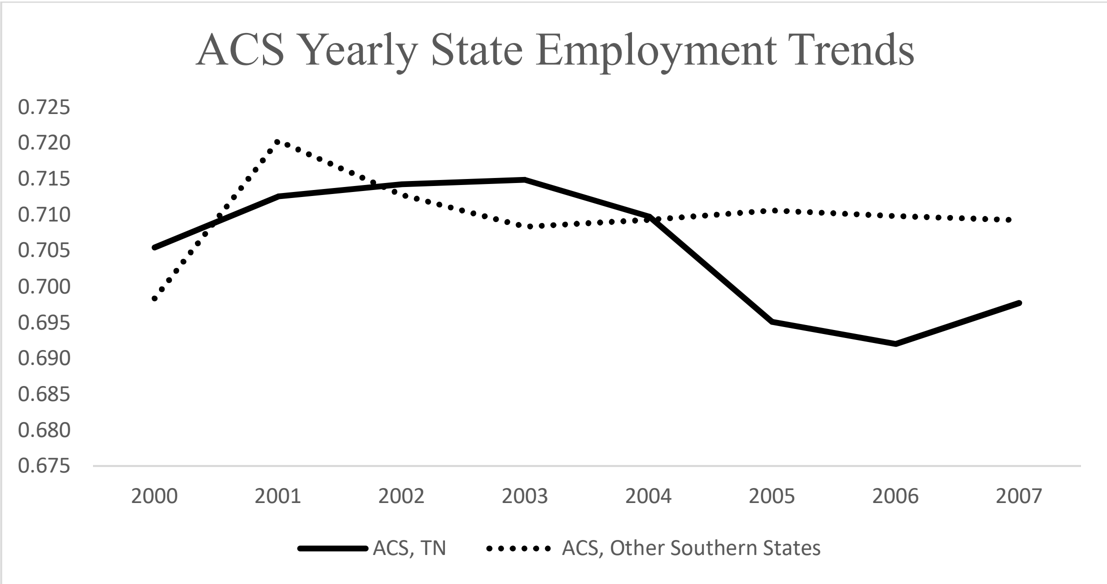
Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the employment share; the X-axis is the year.

Figure 1B: AllCPS Yearly State Employment Trends



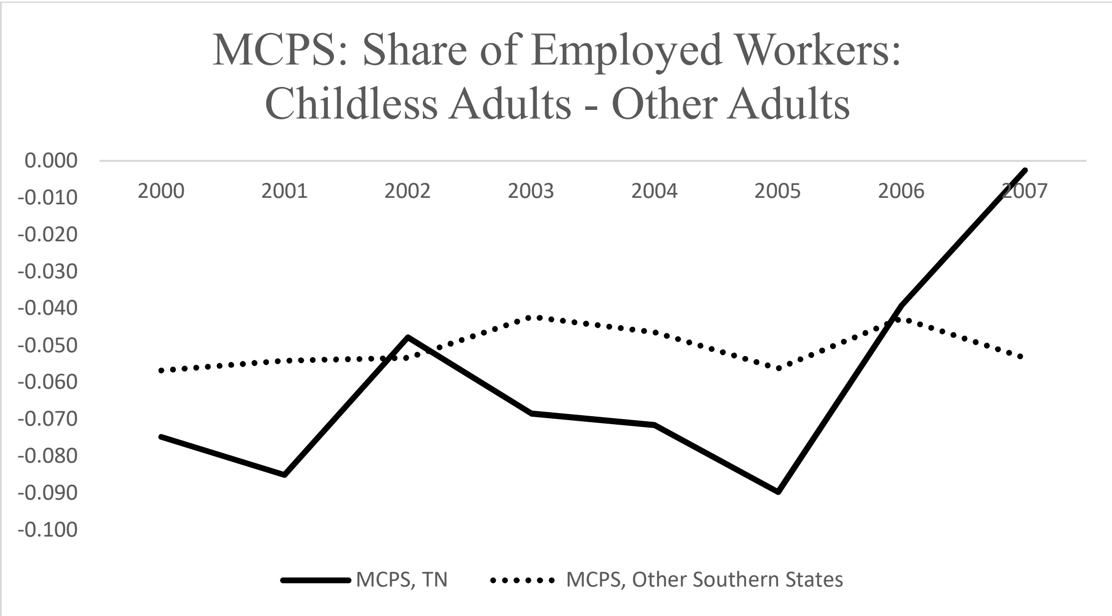
Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the employment share; the X-axis is the year.

Figure 1C: ACS Yearly State Employment Trends



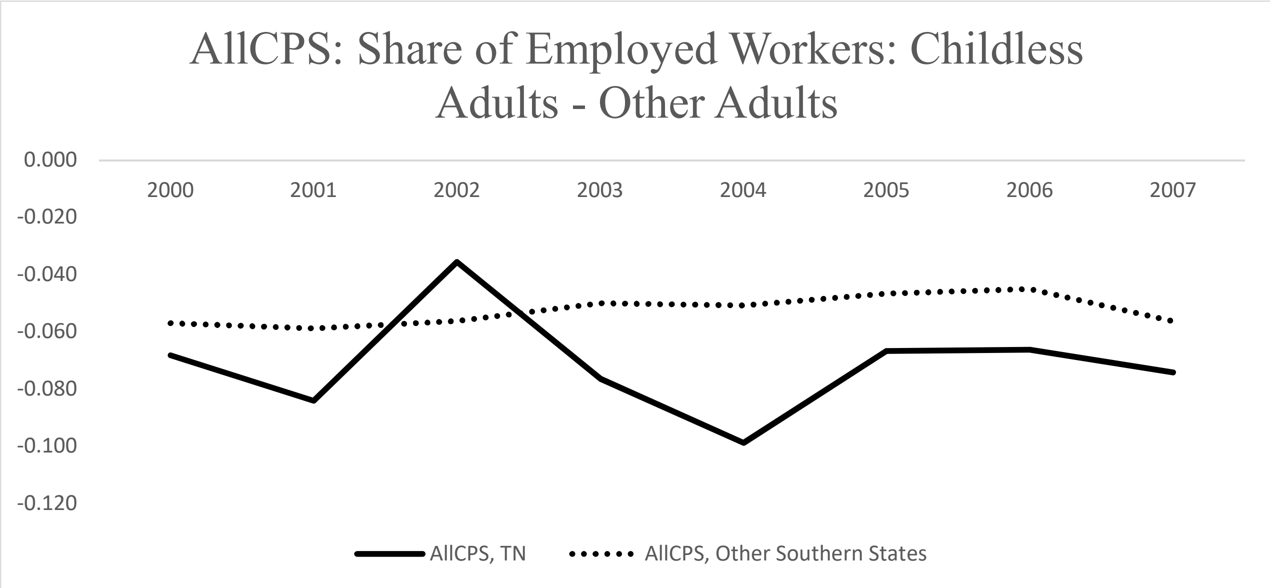
Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the employment share; the X-axis is the year.

Figure 2A: MCPS: Share of Employed Workers; Childless Adults – Other Adults



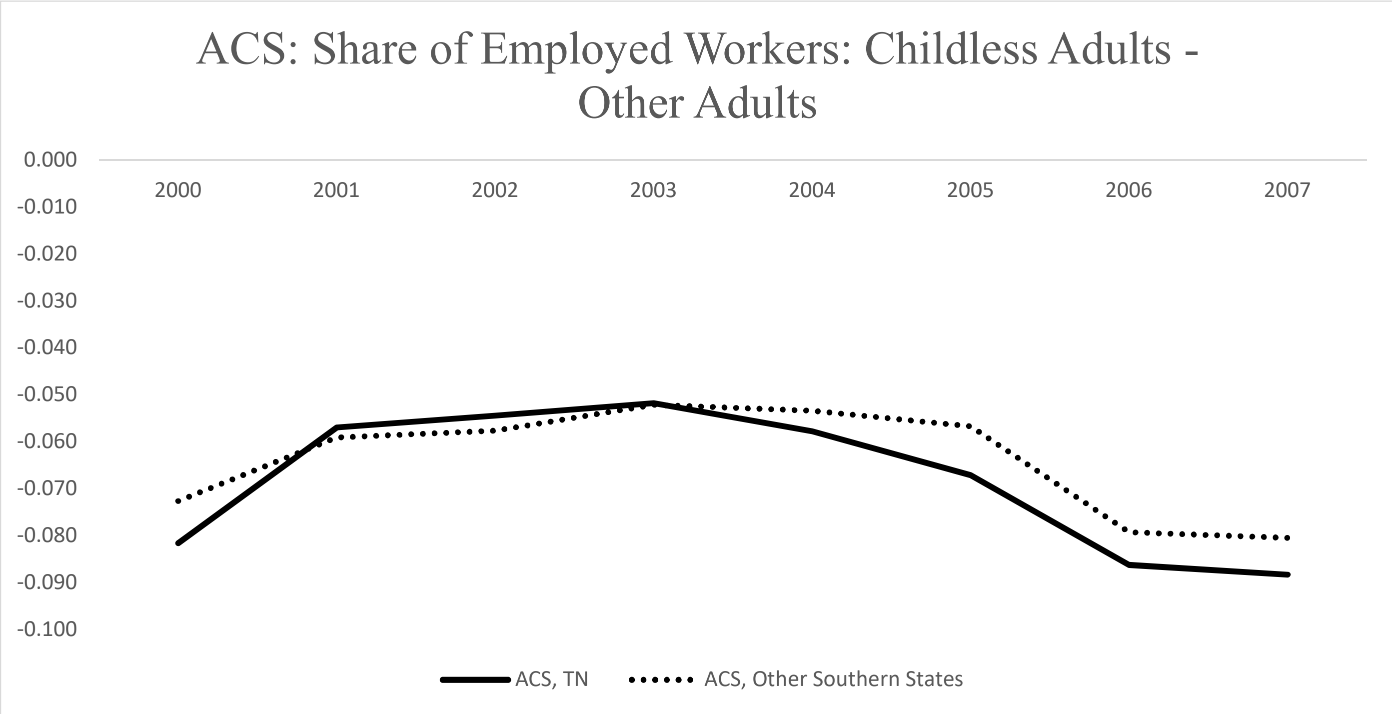
Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the difference in the employment share between those without children under 18 years old vs. those with children under 18 years old. The X-axis is the year.

Figure 2B: AllCPS: Share of Employed Workers; Childless Adults – Other Adults



Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the difference in the employment share between those without children under 18 years old vs. those with children under 18 years old. The X-axis is the year.

Figure 2C: ACS: Share of Employed Workers; Childless Adults – Other Adults



Notes: The years used are every year from 2000 to 2007. The sample of workers is the same sample as the one used in Table 1. The Y-axis is the difference in the employment share between those without children under 18 years old vs. those with children under 18 years old. The X-axis is the year.

Appendix Table 1: TennCare Effect on Employment by Education Attainment

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Table 1 Estimates					
Point Estimate	0.025**	0.022**	0.020*	0.013***	-0.011***
Standard. Error	(0.011)	(0.010)	(0.011)	(0.004)	(0.004)
High School					
Point Estimate	--	0.009	0.016	0.029***	-0.022**
Standard. Error	--	(0.028)	(0.033)	(0.011)	(0.009)
High School or more					
Point Estimate	--	0.025**	0.022*	0.012***	-0.008**
Standard. Error	--	(0.011)	(0.012)	(0.004)	(0.004)
N	136	136	136	136	136
Panel B: Triple Difference					
Table 1 Estimates					
Point Estimate	0.046**	0.055***	0.070***	0.014*	-0.002
Standard. Error	(0.020)	(0.020)	(0.024)	(0.008)	(0.006)
Less Than High School					
Point Estimate	0.125**	0.193***	0.092	0.044**	0.002
Standard. Error	(0.054)	(0.059)	(0.072)	(0.022)	(0.019)
High School or more					
Point Estimate	0.034	0.038*	0.069***	0.011	-0.001
Standard. Error	(0.023)	(0.021)	(0.025)	(0.008)	(0.006)
N	272	272	272	272	272
Microdata Sample: < HS	--	40,843	25,635	313,351	415,638
Microdata Sample: HS or More	--	208,716	141,733	1,744,350	2,620,699

Notes: We use the same years and states as Table 1. We first compute the employment share within each state-year combination for those who had less than a high school education and those who had either a high school degree or higher separately. We then calculate the difference in difference/triple difference estimates for each cohort separately. The standard errors are calculated in the same fashion as Garthwaite et al.

Appendix Table 2: TennCare Effect on Employment by Age

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Table 1 Estimates					
Point Estimate	0.025**	0.022**	0.020*	0.013***	-0.011***
Standard. Error	(0.011)	(0.010)	(0.011)	(0.004)	(0.004)
Age 21-39, Inclusive					
Point Estimate	--	-0.008	-0.019	-0.001	-0.006
Standard. Error	--	(0.015)	(0.018)	(0.006)	(0.005)
Age 40-64, Inclusive					
Point Estimate	--	0.046***	0.050***	0.024***	-0.014***
Standard. Error	--	(0.013)	(0.015)	(0.006)	(0.005)
Panel B: Triple Difference					
Table 1 Estimates					
Point Estimate	0.046**	0.055***	0.070***	0.014*	-0.002
Standard. Error	(0.020)	(0.020)	(0.024)	(0.008)	(0.006)
Age 21-39, Inclusive					
Point Estimate	0.01	0.02	0.054	0.013	-0.007
Standard. Error	(0.031)	(0.029)	(0.036)	(0.011)	(0.009)
Age 40-64, Inclusive					
Point Estimate	0.060**	0.068**	0.058*	0.009	0.007
Standard. Error	(0.028)	(0.027)	(0.033)	(0.010)	(0.008)
N		272	272	272	272
Microdata Sample: Age 21-39		40,843	25,635	313,351	415,638
Microdata Sample: Age 40-64		208,716	141,733	1,744,350	2,620,699

Notes: We use the same years and states as Table 1. We first compute the employment share within each state-year combination for those in the following age brackets separately: 21-39, 40-64. We then calculate the difference in difference/triple difference estimates for each cohort separately. The standard errors are calculated in the same fashion as Garthwaite et al.

Appendix Table 3: Placebo Tests Assuming a 2002 Treatment Year

	GGN	MCPS	BMCPS	ACPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Point Estimate	--	-0.023**	-0.018	-0.007*	-0.001
Standard Error	--	(0.012)	(0.013)	(0.004)	(0.004)
N	102	102	102	102	102
Panel B: Triple Difference					
Point Estimate	--	0.001	-0.014	-0.010	-0.003
Standard Error	--	(0.022)	(0.026)	(0.008)	(0.007)
N	204	204	204	204	204

Notes: We use the same states as Table 1. We use the years 2000 to 2005 instead of 2000 to 2007. All restrictions and calculations are done in the same manner as in Table 1. We treat 2002, 2003, 2004, and 2005 as the treatment years as the placebo test.

Appendix Table 4: Placebo Tests Assuming a 2004 Treatment Year

	GGN	MCPS	BMCPS	ACPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Point Estimate	--	-0.023*	-0.027**	-0.012***	-0.007*
Standard Error	--	(0.012)	(0.013)	(0.004)	(0.004)
N	102	102	102	102	102
Panel B: Triple Difference					
Point Estimate	--	-0.010	-0.016	-0.017**	-0.007
Standard Error	--	(0.023)	(0.025)	(0.008)	(0.008)
N	204	204	204	204	204

Notes: We use the same states as Table 1. We use the years 2000 to 2005 instead of 2000 to 2007. All restrictions and calculations are done in the same manner as in Table 1. We treat 2004 and 2005 as the treatment years as the placebo test.

Appendix Table 5: Placebo Tests Assuming a 2003 Treatment Year – Treatment Effect Varies with Education

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
			Panel A: Difference-in-Difference		
			Less Than High School		
Point Estimate	--	-0.029	-0.070**	-0.025**	0.002
Standard. Error	--	(0.029)	(0.033)	(0.010)	(0.011)
			High School or more		
Point Estimate	--	-0.033***	-0.021*	-0.014***	-0.003
Standard. Error	--	(0.012)	(0.013)	(0.004)	(0.004)
			Panel B: Triple Difference		
			Less Than High School		
Point Estimate	--	-0.015	-0.059	0.025	-0.002
Standard. Error	--	(0.054)	(0.062)	(0.020)	(0.025)
			High School or more		
Point Estimate	--	-0.013	-0.026	-0.027***	-0.008
Standard. Error	--	(0.022)	(0.025)	(0.008)	(0.007)

Notes: We use the same states as Table 1. The outcomes and comparisons are the same as Table 3. We use 2000 to 2005 instead of 2000 to 2007. We use 2003, 2004, and 2005 as the treatment years in the difference in difference analysis.

Appendix Table 6: Placebo Tests Assuming a 2003 Treatment Year – Treatment Effect Varies with Age

	GGN	MCPS	BMCPS	AllCPS	ACS
	(1)	(2)	(3)	(4)	(5)
Panel A: Difference-in-Difference					
Age 21-39, Inclusive					
Point Estimate	--	-0.022	-0.006	0.001	-0.002
Standard. Error	--	(0.015)	(0.017)	(0.005)	(0.005)
Age 40-64, Inclusive					
Point Estimate	--	-0.033**	-0.038**	-0.028***	-0.001
Standard. Error	--	(0.015)	(0.017)	(0.006)	(0.005)
Panel B: Triple Difference					
Age 21-39, Inclusive					
Point Estimate	--	-0.009	-0.041	-0.018*	-0.001
Standard. Error	--	(0.031)	(0.034)	(0.010)	(0.011)
Age 40-64, Inclusive					
Point Estimate	--	0.000	-0.008	-0.009	-0.012
Standard. Error	--	(0.028)	(0.034)	(0.011)	(0.010)

Notes: We use the same states as Table 1. The outcomes and comparisons are the same as Table 4. We use 2000 to 2005 instead of 2000 to 2007. We use 2003, 2004, and 2005 as the treatment years in the difference in difference analysis.