The Efficacy of Hiring Credits in Distressed Areas

Jorge Pérez Pérez[*] and Michael Suher[†]

July 2022

Abstract: We analyze the efficacy of hiring tax credits, particularly in distressed labor markets. These programs have proven hard to assess, as their introduction at the state level tends to be endogenous to local conditions and prospects. We conduct an empirical study of a hiring tax credit program implemented in North Carolina in the mid-1990s, which has a quasi-experimental design. Specifically, the state's 100 counties are ranked each year by a formula trying to capture their economic distress level. The generosity of the tax credits has discrete jumps at various ranking thresholds allowing for the use of regression discontinuity methods. Our estimates show sizable and robust effects on employment - a $9,000 credit leads to a 4% increase in employment or 1.2 percentage point increase in the employment to population ratio in the counties where the hiring credits were available. The attendant reduction in the unemployment rate is around 0.3% percentage points.

1 Introduction

Hiring tax credits are a commonly used tool to address both short-run downturns and longer-run economic distress. The empirical evaluation of these policies is difficult as their enactment is endogenous to expected economic prospects or local labor market distress. The direction of the bias is also not clear. Poor economic performance may swamp estimates of program impacts even if they are positive. Alternatively, natural mean reversion — in areas that recently experienced adverse shocks — could be incorrectly attributed to the policy intervention.

Earlier evidence had questioned hiring credit efficacy, often finding zero impacts on employment (*e.g.* Neumark and Kolko (2010); Behaghel, Lorenceau, and Quantin (2015)). In contrast, Neumark and Grijalva (2017) find positive effects of state-level programs on employment during recessionary periods in the United States. Similarly, Cahuc, Carcillo, and Barbanchon (2019) sees benefits during the Great Recession of a hiring credit program targeted at small firms and low-wage workers in France. We build on this work by evaluating place-based hiring credits aimed at addressing regional disparities in a setting that permits causal estimates. We find significant reductions in unemployment rates and increases in employment in the distressed areas targeted by the hiring program.

We examine a program in the state of North Carolina that assigns county-level hiring credit eligibility based on an economic distress rank. The county rankings include thresholds at which credit size jumps discontinuously. Because program eligibility changes yearly, we implement a dynamic regression discontinuity (RD) design following Cellini, Ferreira, and Rothstein (2010). We complement these findings with local nonparametric estimates that provide consistent evidence. We estimate an intent-to-treat effect of a 4% increase in employment from a $9,000 credit. This effect corresponds to a 1.2 percentage point increase in the employment to population ratio, which we document is driven by increased hiring, as separation rates also rise. In counties where the hiring credits were available, unemployment rates decreased by 0.3 percentage points. Further, employment impacts concentrate in the manufacturing and wholesale trade sectors that were the program's primary target. Finally, we produce difference-in-differences estimates to better nest the

paper with previous attempts at evaluating hiring credits that rely on that technique and to allow us to characterize the direction and extent of endogeneity bias. Difference-in-differences estimations would have failed to find an employment impact, indicating the importance of accounting for time-varying unobservables.

Economists often viewed hiring credits — long a common policy tool — dimly because of concerns around churning, stigmatization, and inefficient distortion of the location of economic activity (Grijalva and Neumark, 2014; Kline and Moretti, 2014). The theoretical case for hiring credits is straightforward: They induce employment increases by reducing the effective wage. When put into practice, they have proved costly or simply ineffective. The first issue is the high baseline rate of hiring and separations. When these incentives are made broadly available, this natural churn means many credits will be claimed for hiring that would have occurred even in the absence of the program. In a meta-analysis, Bartik (2001) finds the so-called wastage rate — the share of credits claimed for hires that would have occurred in the absence of the credits — can approach 90%, effectively inflating the cost per program-induced hire by a factor of 10. For example, in a study of a hiring tax credit program in Georgia, Faulk (2002) finds a wastage rate of 75% but notes that even after adjusting for wastage rates of this magnitude, costs per hire can still compare favorably to discretionary incentive packages granted to a specific firm. Administrative mechanisms exist for addressing the wastage issue so that the credits only reward net job creation. However, the compliance burdens are often so high as to make the credits unattractive to firms. A second issue that appears to undercut hiring credit effectiveness in practice is stigmatization. By targeting specific disadvantaged groups for credit eligibility, these programs can signal low productivity to employers that more than offsets the credit incentive.

That the North Carolina program successfully induced employment increases in a relatively cost-effective manner may be attributable to the economic context around the credit's deployment, and specific features of the program that limited wastage and avoided stigmatization. Neumark (2013) highlights the potential outsized benefit of policies, like hiring credits, aimed at stimulating labor demand during a recession, where downward wage rigidities will mute the impact of labor

3

supply policies. Ku, Schönberg, and Schreiner (2020) find evidence consistent with this framework when a European Union-mandated tax rate harmonization in Norway led to payroll tax increases for large employers in previously subsidized remote areas. The North Carolina program targeted credits to long-run distressed areas with significant labor market slack, less wage pass-through, and a lower baseline hiring rate, reducing wastage. We show that treated counties actually had zero absolute employment growth under the program in our setting, but this contrasted with substantial employment *declines* in the most comparable control counties. Churn and administrative burdens were suppressed through a four-year payout structure for the credits, only requiring that firms demonstrate their employment levels remained elevated each year after the credit's generation. Finally, the credits did not require that hires to be from a specific pool of people, avoiding stigmatizing potential workers. We do, though, find suggestive evidence, in the lack of firm and population entry, that labeling entire counties as distressed may have stigmatized the locales as low productivity.

More recently, a theoretical case for the efficacy of hiring credits in some circumstances has begun developing as spatial inequalities and increasing home attachment have become better understood. Kline and Moretti (2013) augment a spatial equilibrium model with persistent long-run differences in unemployment rates across areas, as is empirically observed. Firms in high unemployment and low productivity areas post too few vacancies due to excessive hiring costs, providing a rationale for subsidizing hiring in distressed areas. Amior and Manning (2018) find that distressed areas experience serially correlated negative demand shocks that lead to swift and continual employment outflows. Population leaves as well, but not as fast as employment, leading unemployment rates to remain elevated. Coate and Mangum (2019) document an increase in workers' home attachment across U.S. labor markets, which translates into declining labor mobility and more persistent unemployment differences across places. Zabek (2019) models how residents' local ties and outsiders' limited interest in moving in imply that place-based policies targeting depressed areas can transfer income with limited geographic distortion.

Our paper contributes most directly to a literature that has sought to empirically measure if, to

what degree, and in what circumstances hiring credits are effective. Credible research designs are needed to address the inherent endogeneity of hiring program adoption and eligibility with current and expected economic performance. Neumark and Grijalva (2017) use cross-state variation in hiring credit programs to estimate their impact, along with counterfactual employment trends based on state industrial composition. They find no impacts on employment growth in general but positive effects during recessionary periods and when programs incorporate recapture provisions. Cahuc et al. (2019) use difference-in-differences and instrumental variable (IV) strategies to evaluate a hiring credit program introduced in France during the Great Recession restricted to small firms and low-wage workers. They find significant impacts on employment at eligible firms. They also find the program to have been particularly cost-effective, though they demonstrate through simulations that this cost-effectiveness depended on the program being both temporary and unanticipated. Sestito and Viviano (2018) evaluate hiring credits in Italy during 2015 that targeted workers without permanent contracts. They find increased hiring for this group of workers in the aftermath of the policy relative to non-eligible workers. Huttunen, Pirttilä, and Uusitalo (2013) find a payroll tax reduction targeting old, low-wage workers in Finland did not affect employment, implying inelastic labor demand for this specific group of workers.

Unlike these interventions, the program we investigate targets specific distressed labor markets. Previous work on place-based hiring subsidies has focused on neighborhood-level programs in urban areas. Most have failed to detect beneficial effects, though there are some exceptions (Papke, 1994; Neumark, 2020). The federal Empowerment Zone (EZ) program in the United States, which includes hiring credits, was studied by Busso, Gregory, and Kline (2013), who identify program impacts using areas with rejected EZ applications or later-round EZ designations as control areas. Freedman (2013) studies the similarly conceived Enterprise Zone program using an RD design approach made possible by a specific program implementation in Texas relying on neighborhood-level poverty rate cutoffs. Both studies find positive employment effects for residents of neighborhoods selected for the programs. In contrast, Neumark and Young (2019) use propensity score matching to evaluate both state Enterprise Zone and federal EZ programs. They find evidence of reduced

unemployment rates but no employment increases or poverty reductions with their preferred empirical strategy. Similarly, in a study of a French EZ program, Gobillon, Magnac, and Selod (2012) find moderate though temporary reductions in unemployment.

Chirinko and Wilson (2016) analyze fiscal foresight, another potential source of bias in evaluations of hiring credits, wherein pre-announced programs could lead firms to initially depress hiring and then ramp up once the credits become available. Exploiting cross-state variation in hiring credit adoption, they report positive impacts on employment at a lag of two to three years, consistent with our findings. They also find pre-program dips, which can upwardly bias estimates of program effects by 33%. Our reliance on annual rather than monthly employment levels should help alleviate this bias.

We describe the mechanics of North Carolina's hiring tax credits in section 2. Section 3 summarizes our data sources and gives an overview of the labor market during our sample period. We describe the estimation strategy in section 4. Section 5 presents the estimation results. Lastly, section 6 offers some discussion of our findings.

2   North Carolina's Hiring Tax Credit Programs

In the mid-1980s, North Carolina government officials were concerned with the divergence in economic fortunes among the state's 100 counties. A tax incentive program began in 1988 to address the situation in the least economically robust counties. The North Carolina Department of Commerce (NC Commerce) ranked counties by the level of economic distress each year, from 1 to 100, using a legislatively defined formula and inputs. The state government used the rankings to segment counties into discrete tiers that determined the size of tax credits a county could claim. The program experienced significant revisions in 1996 and 2007 but maintained the basic ranking calculation and tiers framework. The program's final iteration ended in 2014, but NC Commerce continued computing the county rankings (Program Evaluation Division, 2015). Table 1 summarizes the tax credit size histories.

Table 1: Credit Size by Distress Rank

(Nominal Dollars)

Years	**Distress Rank**									
	10	20	30	40	50	60	70	80	90	100
Wave 0: 1988–95		2,800								
Wave 1: 1996–2006	12,500	3,000–4,000				500–1,000				
Wave 2: 2007–13	12,500					5,000			750	

Note: Lower-ranked counties are more economically distressed. Source: North Carolina Department of Commerce.

In this study, we focus on the William S. Lee program that began in 1996, which we denote as Wave 1 in Table 1, as it provides the cleanest quasi-experimental set-up.[1] Unlike its predecessor, the Lee program extended eligibility to firms in all 100 counties but continued to reserve the larger credits for firms in the more distressed counties. Credits of $12,500 were available to the ten most distressed counties designated as Tier 1. Firms in moderately distressed Tier 2 counties could receive credits between $3,000 and $4,000 per new hire, and those in the least distressed Tier 3 counties could receive between $500 and $1,000.[2] Our analysis will focus on comparing Tier 1 and Tier 2 and the average differential credit size of $9,000.[3]

The Lee program specified county rankings based on three inputs: unemployment rates, income per capita, and population growth. The appendix Table A.1 shows an example of this process for one year of the program. Counties are first ranked separately, 1 to 100, on each input. These three sub-rankings are summed and ordered to create the 1 to 100 distress rank for the coming year. The ten lowest-ranked, most distressed counties place in Tier 1. Counties ranked 11 or higher start in Tier 2. Throughout the program, the number of counties designated Tier 1 each year increased because of the addition of low population and high poverty rates as overrides to the distress ranking

---

[1] The appendix provides details about wave 0 and wave 2 of the tax credit program.

[2] Under the official program definitions, wave 1 of the program has five tiers. Tier 1 is the same as our definition. We combine Tiers 2 and 3 and Tiers 4 and 5, as they have similar program intensity as measured by the credit size for which they are eligible.

[3] This differential is the Tier 1 $12,500 minus the average credit of $3,500 for Tier 2. The program's previous iteration is a confound for comparing counties between Tier 2 and Tier 3. Even though the ranking procedure changed somewhat between wave 0 and wave 1, there is still a discontinuous change in the probability of wave 0 program eligibility at the rank-50 threshold separating Tier 2 and Tier 3.

system.[4] By 2006, the final program year, 28 counties were in Tier 1 and eligible for the largest credits.

Because counties were re-ranked every year, treatment status was not always constant, with occasional slippage between tiers, and the legislated expansion of the lowest tier over time. Figure 1 shows a map with the geographic distribution of county tier designations for the first and final years of the Lee program.[5]

Only firms in specific industries were eligible for the program, with the main ones being manufacturing, wholesale trade, warehousing, and those related to data processing. While hiring certain types of workers, such as those currently unemployed, was not required, they had to be new employees (*i.e.*, not intra-firm transfers from another area in the state), working full time, and paid above the county average wage. The workplace county determined the credit size, which did not depend on an employee's county of residence. The credit paid out over four years, with later installments forfeited if the firm reduced its total number of employees.

Beyond hiring tax credits, the Lee program offered other incentives, most notably for investment in machinery and equipment (M&E) and research and development (R&D) expenditures. These additional incentives had softer discontinuities in generosity at tier thresholds. Their benefits flowed primarily to the least distressed Tier 3 counties, allowing us to isolate the impact of hiring credits from these other incentives.[6]

---

[4]The overrides were as follows: A county's first-time designation to Tier 1 would persist through the subsequent year, even if its subsequent ranking would place it in a higher tier. Starting in 2000, counties with a population of fewer than 10,000 and a poverty rate greater than 16% were automatically placed in Tier 1. The population threshold increased to 12,000 in 2002. Additionally, counties with a population of fewer than 50,000 and a poverty rate greater than 18% had their tier reduced by one from what their distress rank would otherwise dictate.

[5]One prominent feature of the map is the clustering of distressed counties in the eastern part of the state. These are the least urbanized counties and tend to have lower income per capita and population growth, directly affecting their rankings. They also have low population levels. Once the tier designation overrides for population and poverty levels were introduced in 2000, they led even more counties in this region to be designated into the most distressed Tier 1, as seen in panel b of Figure 1. An NC Commerce assessment of the program (Fain, 2001) notes that this targeting was intentional in its design. Plant closings in traditional North Carolina industries like textiles, apparel, furniture, and tobacco manufacturing were disproportionately negatively impacting the state's more rural and economically distressed parts.

[6]The R&D credit size did not vary with the tier system. The M&E credit was constant across tiers but applied to any size investment by Tier 1 firms but only investment greater than $100,000 to $200,000 for Tier 2 firms. Over three-quarters of M&E and R&D credit dollars went to firms in the least distressed Tier 3, which we exclude from the analysis. The Lee program also included separate credits for job training and central administrative offices. However, these were minor, with each accounting for less than 2% of total Lee Act credits generated (Fain, 2005). Cerqua and

Figure 1: County Tier Status

(a)

1996

(b)

2006
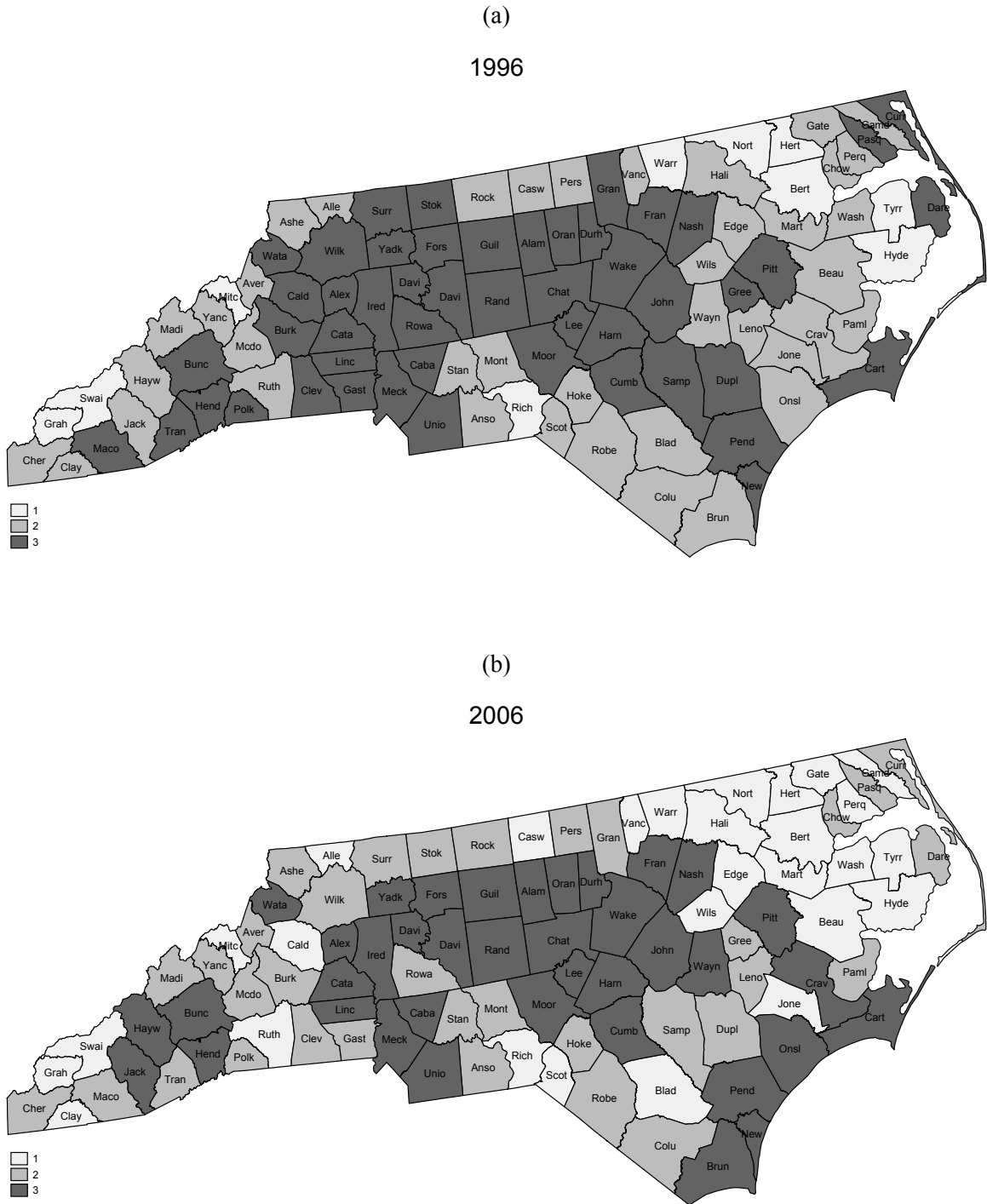
Note: Tier status from NC Commerce for the 100 counties in North Carolina is presented for the first and final years of the William S. Lee program. Tier 1 counties are the most economically distressed.

Figure 2 shows unemployment, income per capita, and population growth for the counties sorted by the distress rank of each input in 1996 and each county's assigned tier. Two facts stand out. First, slight differences in an input variable can lead to large differences in input-specific sub-rankings. Second, counties with very similar inputs can end up in different tiers through the effects of adding up the three sub-rankings and the overrides. Both overall distress rankings and input sub-rankings vary widely over time for each county. On average, a county moves six positions in the overall distress ranking every year, with most of these shifts coming from relative population growth and unemployment rate changes.

In section 4, we demonstrate that lower-ranked counties do have lower population growth, higher poverty, and lower per capita income. However, no evidence indicates discontinuities in pre-treatment conditions at the program thresholds, a fundamental prerequisite for the validity of an RD research design.

Two other potential concerns that could arise with the research design we implement are anticipation effects and manipulation of program participation thresholds. As studied in Chirinko and Wilson (2016), firms may artificially depress current hiring if they anticipate becoming eligible for credits soon, leading to overestimating the program's impact. As mentioned earlier, we use annual rather than monthly data, meaning any distortion in hiring timing induced by the program would need to be over long timescales. Further, the initial enactment of the program occurred in mid-1996 and became effective immediately. Updated tier designations for each future year were not finalized or announced until December of the previous year, limiting the scope for anticipatory hiring delays.

Manipulation of treatment status around the eligibility threshold is always a concern in RD designs. Because the determination of eligibility used county-level data collected by independent official sources and largely depended on relative standing among all other counties, strategic manipulation does not seem possible. One area where some manipulation could have occurred is in

---

Pellegrini (2014) examine the effectiveness of these kinds of capital subsidies in Italy. They exploit discontinuities in subsidy assignments across firms, finding that they effectively boost firm growth. We focus on hiring credits instead of capital subsidies here.

Figure 2: Distress Ranks per Input and Tier Designation

(a)

(b)

(c)
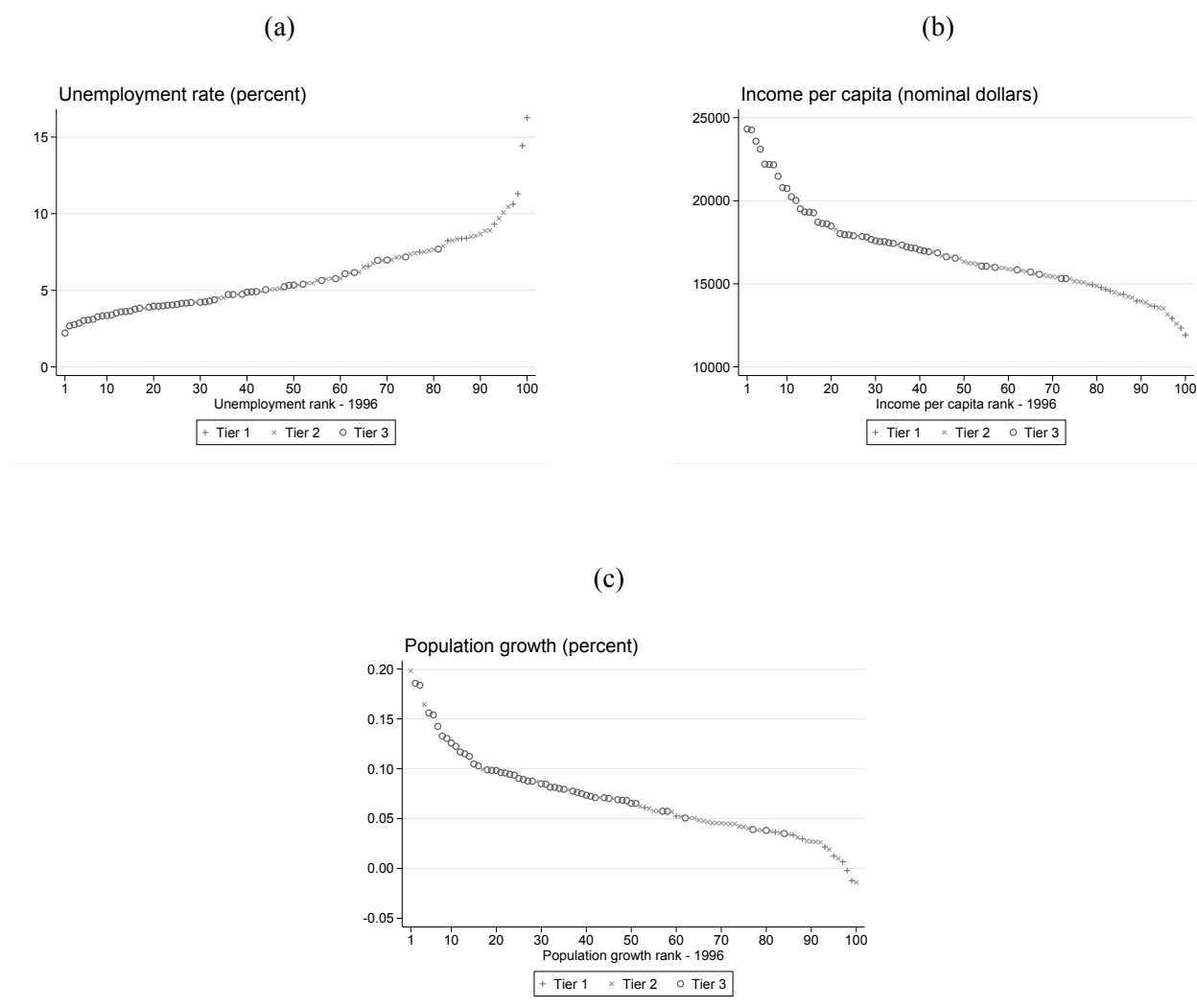
Note: Unemployment data are from the Bureau of Labor Statistics. Population and income per capita data are from the Bureau of Economic Analysis. County-level economic indicators are arrayed by initial distress rank per input at the outset of the first wave of the program. Different symbols denote the different treatment tiers.

revisions to the program, beginning in 2000, which introduced overrides to the tier designation process based on absolute, rather than relative, population and poverty. County lobbying could have influenced the calibration of these overrides, though the choice of round numbers for the thresholds is evidence against any such strategic behavior.

3 Data

Employment, unemployment, and labor force data are from the Bureau of Labor Statistics (BLS). Hiring and separation data come from the Census Bureau's Quarterly Workforce Indicators (QWI). Tier status comes from annual reports issued by the state of North Carolina and archived versions of the NC Commerce website. The distress rankings data is from NC Commerce for some years. For other years, we reconstruct them using data from BLS, population and income data from the Bureau of Economic Analysis, and poverty data from the Census Bureau and the U.S. Department of Agriculture, in conjunction with the rules laid out in the legislation creating and amending the programs. Our sample period runs from 1990 to 2006. Figure 3 shows the overall conditions in North Carolina's economy during the program period. The state's total population is growing throughout the period, highlighting the extent to which the program aimed to address significant within-state regional divergence in economic performance. The state's unemployment rate tracks closely with the national business cycle, and the program span includes the 2001 recession.

Figure 3: Population and Unemployment Rate in North Carolina
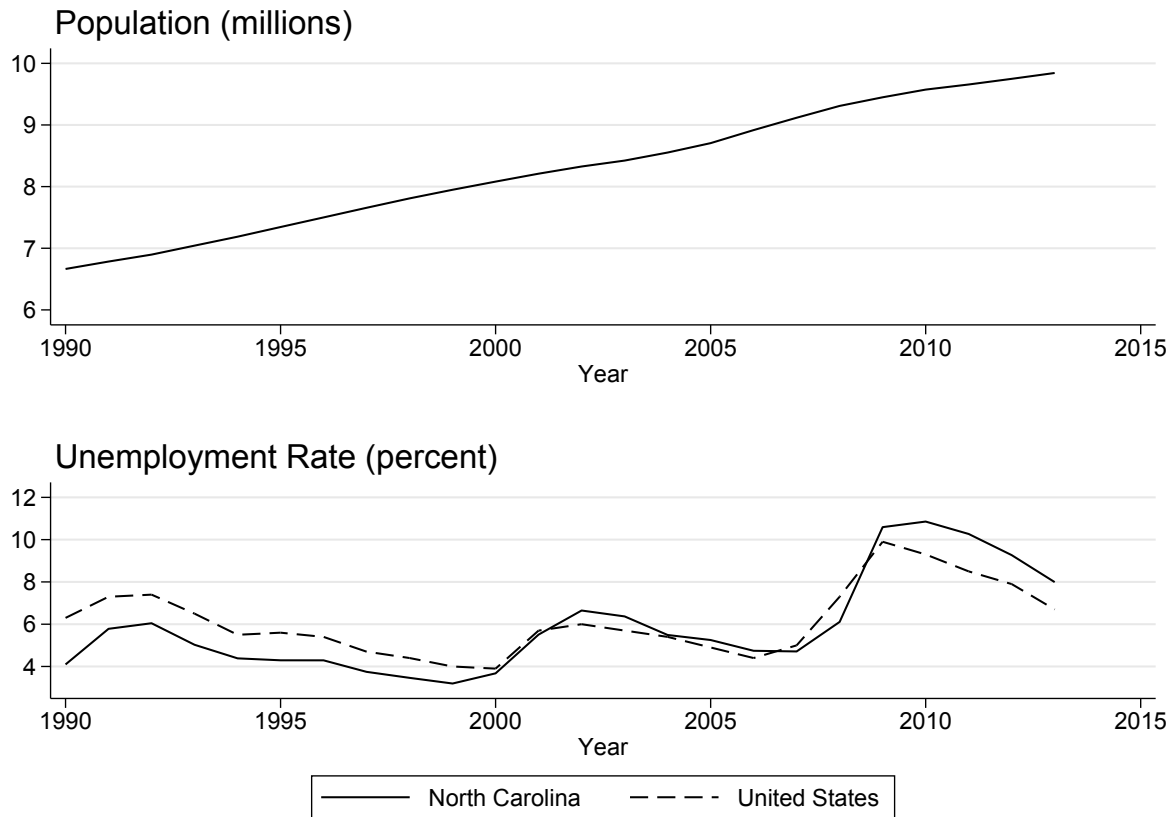
Note: Population data are from the Bureau of Economic Analysis. Unemployment data are from the Bureau of Labor Statistics.

4   Estimating the Effect of Hiring Credits

In this section, we discuss the difficulties of estimating the effect of hiring credits and lay out a strategy to take advantage of the assignment of subsidies based on distress ranks. The institutional setting of the North Carolina hiring credit program calls for the use of a dynamic RD design as well as consideration of multiple treatment assignment variables. North Carolina's 100 counties were assigned annually to three groups defined by the subsidy program tiers. We focus on the wave of the program running from 1996 to 2006. The most distressed counties are in Tier 1 and receive the highest subsidy amount, $12,500. Ten counties were designated Tier 1 in the initial

13

program year, 1996. However, distress rank ties and subsequent amendments to the program that added assignment rules led this number to fluctuate between 10 and 28 in later program years. Tier 1 is our treatment group. Tier 2 counties are the next most distressed and compose our control group. Firms in those counties are eligible for hiring credits ranging from $3,000 to $4,000. The least distressed counties are designated Tier 3 and may apply for hiring credits ranging from $500 to $1,000. We estimate the program's effectiveness by comparing the evolution of employment and unemployment across counties in Tiers 1 and 2. To avoid comparing significantly different counties, we exclude counties always designated as Tier 3. These contain major cities that may have very different dynamics compared with small distressed counties. The average subsidy for Tier 2 counties is $3,500 compared with the $12,500 in Tier 1 counties, so our program effect estimates are for the $9,000 difference.

If the economic distress ranking were completely random, counties would be assigned subsidy amounts randomly, and we could compare counties across tiers. In practice, the distress rank correlates with economic variables. Suppose counties assigned to Tier 1 have worse unobservables, which implies different trajectories of employment and unemployment even in the absence of the program. In that case, estimates from a difference-in-differences approach will be biased. We address this possible bias by exploiting the discontinuities in tier assignments based on the economic distress rank.

The validity of the RD design assumes that county characteristics vary continuously around the tier threshold at distress rank 10 and that only the subsidy amounts vary discontinuously. Figure 4, which shows the relationship between the levels and changes of the outcome variables and the economic distress rank before the beginning of the program, supports this assumption. As expected, there is an overall negative relationship between economic outcomes and economic distress. Unemployment is higher for more distressed counties, while log employment is lower in these counties. However, this relationship is smooth across the tier cutoffs.[7]

---

[7]We cannot reject the null of no break at rank 10 in this relationship for all indicators except log employment, which is expected given that the ten most distressed counties are smaller in population. A potential concern would be different employment trajectories in absence of treatment for these smaller counties. We do not find evidence of such differential growth before 1996, though, because we cannot reject that the relationship between the change in

Figure 4: Relationship between the Distress Rank and Program Outcomes Before the Beginning

of the Program

(a)

Log Employment: 1996

P-value for break test at rank 10 =   0.03

(b)

Employment Growth: 1993–96

P-value for break test at rank 10 =   0.56

(c)

Employment to Population: 1996

P-value for break test at rank 10 =   0.18

(d)

Change in Employment to Population: 1993–96

P-value for break test at rank 10 =   0.88

(e)

Unemployment Rate: 1996

P-value for break test at rank 10 =   0.54

(f)

Change in Unemployment Rate: 1993–96

P-value for break test at rank 10 =   0.74

Note: Unemployment and employment data are from the Bureau of Labor Statistics. Population data are from the Bureau of Economic Analysis. County-level economic indicators are arrayed by initial distress rank at the outset of the first wave of the program. Vertical lines denote the thresholds where credit size jumps discontinuously. The reported p-value comes from a linear regression of each economic indicator on the distress rank and a dummy for a break at 10, restricted to the 50 most distressed counties.

Distress rankings and thus tier assignments are recomputed before each year of the program, meaning treatment patterns—the frequency and timing with which a given county's firms are eligible for the largest subsidies—vary. To account for this facet of the program, we follow Cellini et al. (2010) in implementing a dynamic RD design. This estimation strategy exploits the yearly changes in tier assignments and allows for contemporaneous and lagged effects of the program.

Our baseline specification is

$$Y_{ctk} = \beta_0 + \gamma_c + \gamma_t + \gamma_k + \theta_k tier1_{c,t-k} + \nu_k f(rank_{c,t-k}) + X_{c,t-k-1}\beta_k + \varepsilon_{ctk}. \tag{1}$$

Here, Y_{ctk} is the outcome of interest for county c at time t measured k years after treatment designation. γ_c, γ_t, and γ_k are fixed effects for the county, year, and time since treatment designation. The variable $tier1_{c,t-k}$ is a dummy variable equal to 1 whenever a county is in Tier 1. The θ_k coefficients measure the program impact at various lags. The function $f(rank_{c,t-k})$ depends on the county ranking at time $t - k$. The variables $X_{c,t-k-1}$ are control variables measured at $k - 1$, before the program effects start. When included, the ranking and controls coefficients can vary with time since treatment assignment k but do not vary by calendar year t.

Allowing for lagged effects is essential, as the hiring subsidy programs may take a few years to gain traction and have a noticeable impact on employment (Neumark and Grijalva, 2017). We can assess how the program's impact changes over time by looking at the differences in these coefficients. We estimate this specification using only counties in Tiers 1 and 2, from 1996 to 2006.

The RD approach allows us to obtain unbiased estimates of the program's effect as long as the conditional expectation of the unobservables that enter ε_{ctk} in (1) with regard to the county ranking varies smoothly across the Tier 1 cutoff. Additionally, these estimates will address bias from mean

---

employment from 1993 to 1996 and the distress rank is continuous at the rank ten cutoff. We do not find discontinuities in any economic outcome on a six ranks window around the cutoff, using a local randomization test (such as the ones used in section 5.1). In Appendix Figure B.1, we show the absence of discontinuities in the college share and the target employment share before the beginning of the program. In Appendix Figure B.5, we also show that the relationship between the distress rank inputs and the variables that determine overrides to tier assignment is smooth across tier cutoffs.

reversion if the mean-reverting component of ε_{ctk} arising from transitory shocks does not change discretely across the cutoff. The estimation strategy addresses mean reversion bias even though Tier 1 assignment depends indirectly on pre-assignment outcomes through the ranking function (Chay, McEwan, and Urquiola, 2005).

There are some issues with implementing this specification in our setting. The first issue is that Tier 1 assignment did not entirely depend on the economic distress rank. Counties could not be re-designated out of Tier 1 because of an improved distress rank until after two years. From 2000 onward, high-poverty- and low-population-based rules enter as overrides to the formula for tier assignment. Wong, Steiner, and Cook (2013) propose and assess methods for dealing with multiple assignment variables in an RD framework. They recommend excluding units assigned based on additional rules and estimating equation (1) as a sharp discontinuity design using only counties assigned based on the running variable being considered—in this case, the distress rank. Another approach is to classify the counties that change tiers because of these overrides as "defiers" and instrument Tier 1 status with Tier 1 assignment based on the distress rank, as in a fuzzy discontinuity design. We focus on the recommended sharp RD strategy in section 5 but also provide estimates using the alternative fuzzy RD approach in the appendix.

The second issue is the sample size available to estimate each year's program effect. For the comparison between Tier 1 and Tier 2, we have only 70 counties available.[8] This sample size limits our ability to estimate a large number of parameters or implement nonparametric estimators. We reduce the number of coefficients to estimate by assuming constant treatment effects and a linear control function with the same slope on either side of the cutoff that is consistent with data before the beginning of the program.

If the program's effect is constant over time, we can take advantage of its repeated execution. Our constant treatment effect assumption is that the program's effect depends only on the number of years that have passed since the program takes place. We generate a new set of observations for out-

---

[8]Tier 1 has counties ranked ten and below, and Tier 2 has counties ranked between 10 and 50. More than 50 counties contribute some observations to the sample because when rankings are recomputed each year, some formerly Tier 3 counties fall into a lower tier. Thirty-two different counties receive treatment in at least one cohort, and fifty-nine different counties are in the control group in at least one cohort.

comes stretching from two years before to four years after treatment designation for each program year and county. We pool together these spans of observations to estimate a single panel regression. So for a given county, there are repeated, overlapping observation windows for [1993,1999], [1994,2000], [1995,2001], etc. for program designation rounds taking place in 1995, 1996, and 1997, etc., respectively. We account for the multiple appearances of a given county-by-year outcome by clustering standard errors by county.

Our additional assumption concerns the functional form of $f(rank_{c,t-k})$. Figure 4 suggests a linear conditional expectation function of changes in the outcomes given a distress rank. Moreover, the functional form of this relationship does not seem to change at the cutoff threshold. Therefore, we assume that $f(rank_{c,t})$ is linear and remains constant on either side of the assignment cutoff.[9] We also try including the ranking input variables themselves—unemployment rate, income per capita, and population growth—as controls.

To address concerns about this functional form assumption and extrapolation far away from the cutoffs, we also calculate local estimates that only use variation near the cutoff. We pool changes in outcomes following each year of the program and use information criteria to select the control function, which often suggests a simple comparison of the means of these changes across tiers (Lee and Lemieux, 2010). We conduct hypothesis tests on these local estimates using randomization inference (Cattaneo, Frandsen, and Titiunik, 2015; Cattaneo, Titiunik, and Vazquez-Bare, 2016).

We also experiment with how we account for the dynamic nature of the subsidy program. Consider the program's effect two years after its enactment in year t. In the year $t+1$, the county would receive the contemporaneous and lagged effect of the program. If the county receives the subsidy in year $t+1$ as well, by year $t+2$ it would experience lagged effects of the program in t and $t+1$ together. Moreover, receiving the program in t may have altered the probability of receiving it at $t+1$. We include indicators for prior treatment status but not for subsequent treatment. Cellini et al. (2010) show that in this setting, the estimated effects have an "Intention to Treat" (ITT) interpretation, where employment outcomes are not only affected by the receipt of the subsidy but also

---

[9]The choice of a linear control function is also informed by Gelman and Imbens (2019), who show that higher order polynomials of the running variable can induce bias in the discontinuity estimates.

by changes in the probability of receiving the subsidy in the future. Importantly, all comparisons keep treatment and control groups constant by program year cohort.

Cellini et al. (2010) also developed a "Treatment on the Treated" (TOT) estimator that accounts for the indirect impact of initial treatment on the probability of treatment in future years.[10] We apply their method with the following regression:

$$Y_{ct} = \beta_0 + \gamma_c + \gamma_t + \sum_{k=0}^{K} (\alpha_k m_{c,t-k} + \theta_k tier1_{c,t-k} m_{c,t-k} + \nu_k m_{c,t-k} f(rank_{c,t-k})) + X_{c,t-K-1}\beta + \varepsilon_{ct}.$$
(2)

Here, Y_{ct} is the outcome of interest for county c at time t. The coefficients γ_c and γ_t are fixed effects for county and year. The variable $tier1_{c,t-k}$ is a dummy variable equal to 1 whenever a county is assigned to Tier 1. The variable $m_{c,t-k}$ is a dummy for being in the 50 most distressed counties at time $t-k$, so the county was either in Tier 1 or Tier 2 at the time. The θ_k coefficients measure the program impact at various lags. The coefficient on the ranking is allowed to vary with time since treatment assignment k. We interact treatment assignment with the $m_{c,t-k}$ dummies to use variation only from the most distressed counties every year. By including the history of treatment assignment up to each outcome observation, θ_k will be a TOT estimate. Including this history will isolate the impact of receiving treatment k years ago and not in subsequent years.

How this TOT estimate would compare to the ITT estimate is *a priori* uncertain in this setting because ITT estimates combine two veins of future treatment designation with the direct impact of initial treatment. First is a potential dampening from some control group members getting credit eligibility in subsequent years. Second is amplification from any re-designation for credit eligibility of treatment group members in subsequent years.

---

[10]The ITT and TOT terminology for the dynamic RD design is from Cellini et al. (2010). It should not be confused with the ITT and TOT terminology for randomized experiments with partial compliance. There is full compliance with the hiring credits program at the county level when accounting for the overrides.

5 Results

We now turn to the RD estimates. The left panels of Figure 5 portray graphical evidence. We array counties by the initial distress rank relative to the threshold where credit size increases. A linear fit in county rank is included, constrained to have the same slope on either side of the threshold. Given the previous evidence that program effects appear only with a lag, we focus on three-year differences. The figure pools outcomes for counties entering the program at different times. The outcomes correlate weakly with the distress rank, with the assumption of a linear relationship with rank appearing reasonable.[11]
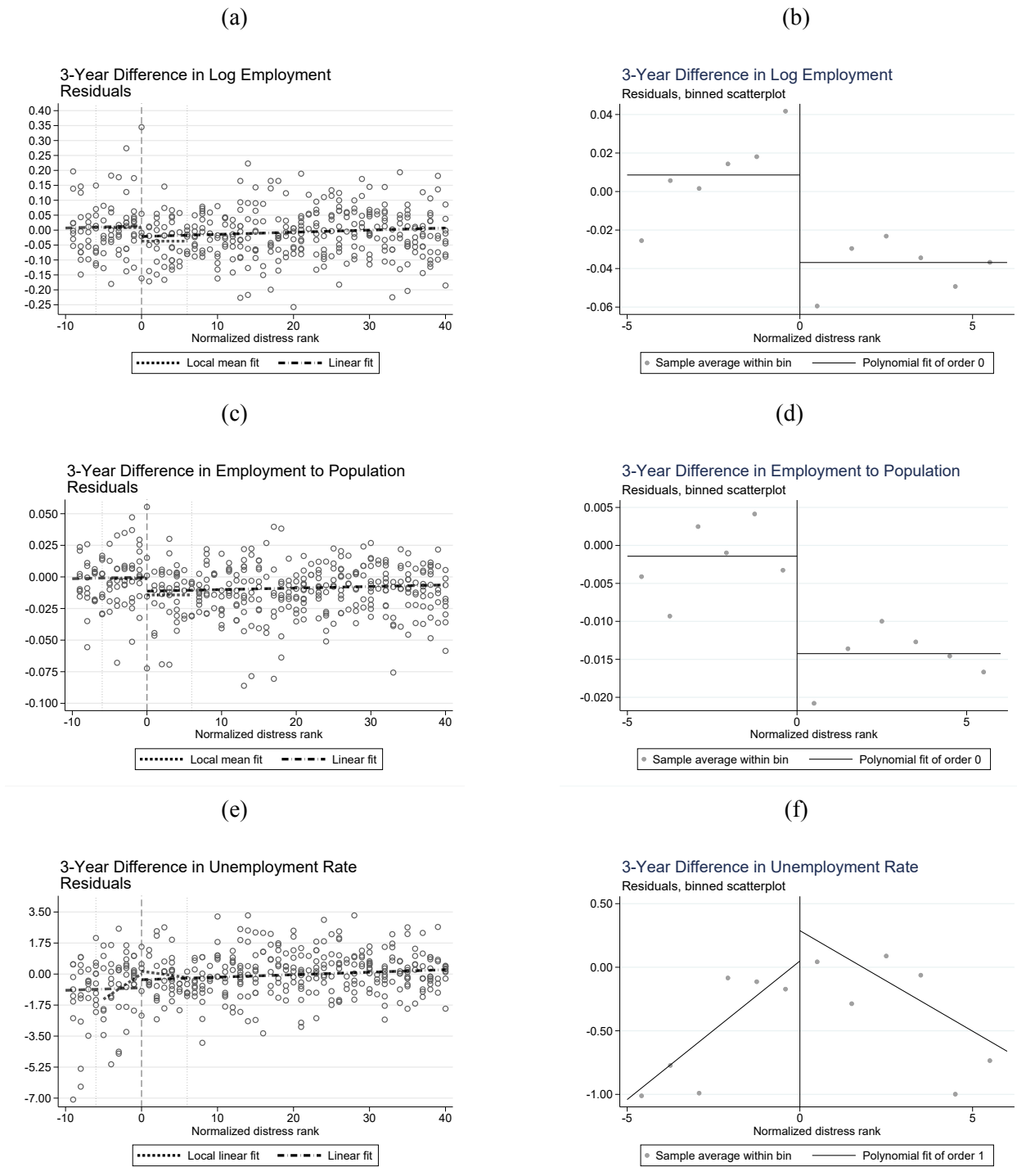
Table 2 presents the dynamic ITT estimates from equation (1). Rows 1 and 2 show estimates for log employment, which are progressively increasing for treated counties relative to the counterfactual through three years after treatment designation. The three-year effect with the ranking input variables as additional controls shows an employment increase of 3.6%. The same pattern of estimates emerges for the employment-to-population ratio in the next two rows. The effect three years after treatment is estimated to be around one percentage point. Rows 5 and 6 show that the unemployment rate in Tier 1 counties is lower relative to control counties two years after treatment and continues its relative decline three years after. Relative to control counties, the unemployment rate in treated counties is between 0.5 and 0.7 percentage points lower after three years. The lower estimate is the result of adding additional county-level controls. For reference, throughout the program, unemployment rates averaged 6.6% for the sample overall and 7.9% for the most distressed counties that make up the treatment group.[12]

---

[11]There is an apparent outlier in panel (a) of Figure 5. This data point corresponds to Northampton county in 2005. This large employment effect may be due to the program. Lowe's is the biggest Tier 1 credit recipient, and it has a distribution center in Garysburg in Northampton county. We do not have data on the jobs generated by Lowe's in every county, but aggregate data indicate that Lowe's generated 673 jobs in 2003, 271 in 2004, and 274 in 2005. This change in employment totals accounts for 1,218 direct hires generating credits, close to the entire employment increase shown in figure 5. We show local estimates excluding this data point in appendix table B.4. The results are qualitatively similar.

[12]In appendix table B.1, we allow the linear control function to change slope at the threshold and estimate the treatment effect derivative (TED) discussed in Dong and Lewbel (2015). The TED is the coefficient on the interaction between treatment status and the running variable. The top panel repeats the estimates from table 2 and shows the estimated linear control function. The lower panel repeats these estimations with the interaction term included. The estimated program effect's magnitudes are unchanged by this more flexible approach, but standard errors increase. The TED is never statistically significant, so we assume a constant slope in all parametric estimates and exclude this

Figure 5: Discontinuities in 3-Year Differences of Employment and Unemployment.

(a)

3-Year Difference in Log Employment
Residuals

(b)

3-Year Difference in Log Employment
Residuals, binned scatterplot

(c)

3-Year Difference in Employment to Population
Residuals

(d)

3-Year Difference in Employment to Population
Residuals, binned scatterplot

(e)

3-Year Difference in Unemployment Rate
Residuals

(f)

3-Year Difference in Unemployment Rate
Residuals, binned scatterplot

Note: Data are from the Bureau of Labor Statistics. Three-year differences in outcomes, 1996 to 2006. Panels (a), (c) and (e) show sample means plus residuals of a regression of the differenced outcomes on year dummies. Counties are arrayed by distress rank relative to the threshold. Counties to the left of the threshold are eligible for a larger hiring tax credit. The thicker lines are estimated linear control functions in distress rank. The thinner lines are means (panels a and c) or linear fits (panel e) within a bandwidth of +/- 6 distress ranks. Panels (b), (d) and (f) show binned scatter plots of the residuals within a bandwidth of +/- 6 ranks, adjusted for mass points in the running variable following Calonico, Cattaneo, and Titiunik (2015). The lines show local means (panels b and d) or a linear fit (panel f).

The building in effect size year by year may reflect a lagged effect from firms needing to learn about the credits and the persistence of treatment status. A majority of treated counties receive Tier 1 status for more than three years, and most counties receive three years of treatment conditional on receiving treatment the first time.

Table 2: Regression Discontinuity Intention to Treat Estimates - Main Outcomes

Dependent Variable	1 year later	2 years later	3 years later
Log Employment	0.005	0.013	0.030*
	(0.013)	(0.016)	(0.017)
With Controls	0.006	0.016	0.036**
	(0.013)	(0.015)	(0.017)
Employment to Population	0.002	0.004	0.010*
	(0.004)	(0.005)	(0.005)
With Controls	0.002	0.005	0.012**
	(0.004)	(0.005)	(0.005)
Unemployment Rate	0.075	-0.468	-0.748**
	(0.274)	(0.306)	(0.322)
With Controls	0.188	-0.319	-0.507**
	(0.319)	(0.261)	(0.228)

Clustered standard errors are in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. N= 2,779. Each row comes from a separate estimation of equation (1) and reports the treatment effect estimates θ_k at 1, 2, and 3 years after treatment designation. Standard errors are clustered by county. All rows include fixed effects for year, county, time since tier designation, and prior treatment history, as well as a linear control function in the distress rank. Additional controls are the lagged 3-year averages of the unemployment rate and real income per capita and the population growth since the most recent census, which are the 3 inputs to the distress rank. The estimates correspond to a sharp regression discontinuity (RD) design, which excludes from estimation of the treatment effects any counties designated as Tier 1 by an assignment rule besides the primary one based on distressed rank.

Table 3 implements equation (2) for the dynamic TOT estimates. As with the ITT estimates, we use the approach recommended by Wong et al. (2013) for handling multiple assignment rules. Namely, we implement a sharp RD design and exclude counties assigned by rules besides the treatment group's distress rank. All rows include county-level controls in addition to the running

interaction term. The insignificant TED estimates also suggest that program effects would be invariant to marginal changes in the eligibility threshold.

variable. For employment growth, the treatment effect estimates are positive three years later but statistically insignificant. This lack of statistical significance likely reflects the more demanding nature of this estimation relative to the ITT specification above, as it estimates extra parameters simultaneously. The unemployment results are clearer than the employment results. Three years after, eligibility for the most extensive hiring credits has reduced the unemployment rate in a county by 0.5 percentage points.[13] Compared to the ITT estimates, these TOT estimates show the effect of being eligible for treatment in year zero and then not again in the subsequent three years. Because such a treatment pattern rarely occurs under this program in practice, with most counties retaining eligibility in subsequent years, we consider the TOT estimates less well founded than the ITT approach.

Table 3: Regression Discontinuity TOT Estimates - Main Outcomes

Dependent Variable - Method	1 year later	2 years later	3 years later
Log Employment	-0.033*	-0.005	0.009
	(0.016)	(0.014)	(0.019)
Employment to Population	-0.008*	-0.001	0.007
	(0.004)	(0.004)	(0.005)
Unemployment Rate	0.150	-0.582*	-0.530*
	(0.343)	(0.337)	(0.306)

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. N=770. Each row comes from a separate estimation of equation (2) and reports the treatment effect estimates θ_k at 1, 2, and 3 years after treatment designation. Standard errors are clustered by county. All rows include lags of the unemployment rate, income per capita, and population growth as controls. The estimates correspond to a sharp regression discontinuity (RD) design, which excludes from estimation of the treatment effects any counties designated as Tier 1 by an assignment rule besides the primary one based on distressed rank. TOT is Treatment on the Treated.

---

[13]In appendix table B.3 we use an alternative estimation approach suggested by Wong et al. (2013). We implement a fuzzy RD design, including all counties assigned to Tier 1 by any rule and instrument for treatment using the primary assignment rule based on distress rank. Overall, these IV estimates are larger, showing sizable effects on employment and unemployment three years after treatment. The employment-to-population results display a similar pattern, but the IV estimates do not show significant employment to population decreases one year after the program begins. We prefer the estimates in Table 3 since Wong et al. (2013) show that the OLS estimator has the least bias in their simulations.

5.1 Local Estimates

The RD design assumes a random treatment assignment at the policy threshold conditional on a control function that allows for more distant observations to contribute to the treatment effect estimation. As such, this estimation requires assumptions about the shape of that control function. Cattaneo et al. (2015) propose a nonparametric estimation technique that uses randomization inference in a small neighborhood around the threshold. We implement their approach in Table 4, where we show estimates for the program's effect three years ahead.

We conduct a series of covariate balance tests for progressively larger bandwidths around the threshold to determine the estimation window where the random assignment assumption is deemed most plausible, . We use the lags of income per capita, unemployment rate, share college, and target industry employment share as the covariates. Figure B.2 in the appendix visualizes the tests, which suggest a bandwidth of +/- 6. We then estimate the treatment effect as a difference in means, which we display as overlays on the left panels of Figure 5. When necessary, we include a polynomial in the running variable to account for residual dependence. In the right panels of Figure 5, we show local mean or linear estimates in a bandwidth of +/-6 ranks, along with binned scatter plots to display the underlying conditional mean function, following Calonico et al. (2015).

The local estimation procedure yields an ITT effect, as it incorporates any impact initial treatment has on the probability of treatment in future years. We calculate p-values from repeated random resampling of the counties within the bandwidth to either side of the threshold. Outcomes are pooled across all program years after year effects have been partialled out.[14]

The estimate for log employment in row one of Table 4 implies that receiving the credits increases employment in a county by 4.5% after three years. The estimate for employment-to-

---

[14]Estimating the treatment effect as a difference in means assumes that the outcome is as good as randomly assigned in a small neighborhood around the cutoff. This assumption may not hold if there is leftover dependence between the outcome and the running variable, even when restricted to the small neighborhood. In appendix table B.5, we follow Cattaneo et al. (2015) and Cattaneo et al. (2016) and allow for leftover dependence. On either side of the cutoff, we fit linear or quadratic polynomials. Then, we estimate the treatment effect as the difference between the polynomials at the cutoff. We select the order of the polynomial adjustment using the Akaike information criterion (Lee and Lemieux, 2010). In most cases, the preferred model is no adjustment. The exceptions are log employment one year after treatment and unemployment 2 and 3 years after treatment, when a linear adjustment is preferred.

population implies an increase of 1.3 percentage points.

The estimate for unemployment in row three of Table 4 implies that receiving the credits decreases the unemployment rate in a county by 0.18 percentage points after three years. This effect is smaller and statistically insignificant in contrast with the parametric estimates above. Expanding the window leads to more precise negative 0.5 percentage points estimates, in line with the parametric estimates.[15]

Table 4: Local Estimates - Main Outcomes

Dependent Variable	1 year later	2 years later	3 years later
Log Employment	0.013	0.022	0.045**
	[0.12]	[0.14]	[0.01]
Employment to Population	0.001	0.006	0.013***
	[0.68]	[0.14]	[0.01]
Unemployment Rate	-0.279	-0.129	-0.182
	[0.12]	[0.58]	[0.53]
Observations	118	106	94
Counties	27	26	24
Ever Treated Counties	18	18	16

P-values from randomization inference with 1000 replications in brackets
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. Difference in mean outcomes for treated and control counties in a window of 6 ranks around the treatment cutoff, adjusting for remaining dependence on the running variable. All estimates are unadjusted differences in means except for log employment 1 year after treatment, and unemployment 2 and 3 years after treatment, which include a linear adjustment on the distress rank. The "Observations" row counts all county-year observations. The "Counties" row counts counties ever included in the estimation. The "Ever Treated Counties" counts counties that were included as treated for at least one year.

---

[15]In appendix figure B.3, we gauge the stability of our three-years-later estimates to different bandwidths. Employment and employment-to-population estimates are somewhat larger for smaller windows but remain positive and significant across different bandwidth choices. Unemployment rate estimates are small though stable at smaller bandwidths though they turn positive with larger comparison windows.

5.2 Additional Outcomes

The tax credits were limited to eligible industries, mainly manufacturing, warehousing, whole-sale trade, and data processing.[16] Rows 1 and 2 of Table 5 present RD estimates of equations (1) and (2) and local estimates for aggregate employment within the targeted industries and non-targeted industries separately. The point estimates for log employment are imprecise but show increases in target industry employment and no evidence of increased employment in non-target industries. The employment-to-population estimates are in rows 3 and 4. These estimates account for the significant variability in small counties' employment at the industry level, showing more precise employment increases. Target industry employment-to-population ratio increases are statistically significant and around one percentage point. ITT target employment estimates are much larger than aggregate impacts, which makes sense as they capture a similar level change against a smaller base value. The target employment to population ITT estimates are equal to the aggregate impacts. We should see this similarity if there are no spillovers to non-target industries, given that the population denominator is the same for target and aggregate estimates. Due to the reduction in sample size from the presence of censored data at the county by industry level, we present local estimates for a window of +/- 10 rather than the preferred +/- 6 ranks. The local estimates, which also capture an ITT effect, tell a similar story of no spillovers (positive or negative) on non-target industries.

In Figure 6, we present the effect by each two-digit industry separately, estimating the impact on the ratio of each industry's employment to the overall county population. The three principal target industries are at the top. The effects are small and insignificant for all the sectors except manufacturing, the largest targeted industry. The estimated increase in the industry employment-to-population ratio is about 1.5 percentage points, similar to the ones estimated in Table 5 for all the targeted industries combined. The lack of discernible impact for warehousing may be due to

---

[16]At the inception of the program in 1996, eligible industries, with North American Industry Classification System codes in parentheses, were Manufacturing (31-33), Warehousing (493), Wholesale Trade (42), Research and Development (541710), and Data Processing (Computer Systems Design & Related Services (54151), Software Publishers (511210), Software Reproducing (334611), Data Processing Services (514210), and On-Line Information Services (514191)). Air Courier Services (492110), Central Administrative Office (551114), Electronic Mail Order (454110), and Customer Service Center (561422) were made eligible beginning in 1999.

Table 5: Other Outcomes 3 Years after Treatment

Dependent Variable	Specification		
	ITT	TOT	Local
Log Employment Target Industries	0.067* (0.039)	0.020 (0.040)	0.018 [0.536]
Log Employment Non-Target Industries	0.003 (0.046)	-0.032 (0.041)	0.002 [0.958]
Employment to Population Target Industries	0.012** (0.004)	0.003 (0.003)	0.008*** [0.001]
Employment to Population Non-Target Industries	-0.001 (0.009)	-0.005 (0.009)	-0.002 [0.717]
Log Hires Annual Total	0.062 (0.040)	0.031 (0.033)	0.103*** [0.007]
Log Separations Annual Total	0.043 (0.039)	0.013 (0.038)	0.088** [0.010]
Log Population	-0.008 (0.007)	-0.016 (0.009)	-0.007* [0.063]
Log Establishment Count	-0.006 (0.013)	-0.014 (0.011)	-0.011* [0.087]
Log Average Weekly Wage	0.003 (0.009)	-0.006 (0.010)	-0.004 [0.662]

Clustered standard errors are in parentheses. P-values are in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. All columns report treatment effect estimates 3 years after treatment designation. Standard errors are clustered by county. The intention to treat (ITT) column reports estimates from equation (1). N=2,779 for hires and separations. N=2,700 for industry employment due to censoring of industry level data in some small counties. All rows include fixed effects for year, county, time since tier designation, and prior treatment history, as well as a linear control function in the distress rank. The treatment on the treated (TOT) column reports OLS estimates from equation (2). N=770 for hires and separations. N=748 for industry employment due to censoring of industry level data in some small counties. The Local estimates column reports the difference in mean outcomes for treated and control counties, on a window of 10 ranks around the treatment cutoff. ITT and TOT estimations include as controls lagged three-year averages of the unemployment rate and real income per capita and population growth since the most recent census, which are the 3 inputs to the distress rank.

its combination with a non-targeted industry, transportation. With wholesale trade and warehousing, the wage requirement was another factor potentially limiting tax credit use. Credits could be claimed only for jobs paying above the county average wage, which is more common in manufacturing than in the other two target industries.[17]

As a robustness check, we re-estimate the RD specifications using Census QWI data on hires and separations. The employment gains we attribute to a hiring tax credit program should stem from increased hiring, not decreased separations. Rows 5 and 6 of Table 5 present the RD estimates for hires and separations following a county's eligibility for the program. The coefficients are consistent with employment gains resulting from increased hiring rather than decreased separations. Separations show evidence of an increase as well, which tracks with the empirical observation in the labor literature that growing firms continue to have positive separation rates (Cahuc, Carcillo, and Zylberberg, 2014). Finally, we examine population and establishment counts to assess possible in-migration response and firm entry. These estimates are noisy but show if anything, net out-migration and firm exit following program assignment. This finding is suggestive that the distressed label may have stigmatized these locales as low productivity, at least for non-incumbent firms.[18]

The last row of Table 5 shows estimates of the effect on average weekly wages by county. The hiring credits do not seem to increase local wages, suggesting that they do not induce significant spillover effects or crowd out employment in other industries. The absence of a wage response is consistent with credits boosting hiring in distressed counties where hiring rates were sub-optimally low (Kline and Moretti, 2014).

---

[17]Jolley, Lane, and Paynter (2013) use detailed firm data to document what types of firms claimed hiring credits under the Lee Act for 2006. Out of 1088 hires in the entire state, 606 (55.7%) went to manufacturing, 141 (13%) went to wholesale trade and transportation, and the remainder went to other industries. Firms in the warehousing industry did not claim credits.

[18]Data for 2006 from Jolley et al. (2013) also suggests a limited impact on out-of-state immigration. They document that 83.9 percent of credits claimed in 2006 went to workers with a previous employment history in North Carolina.

Figure 6: Estimates of Effects on Industry Employment-to-County Population 3 Years after
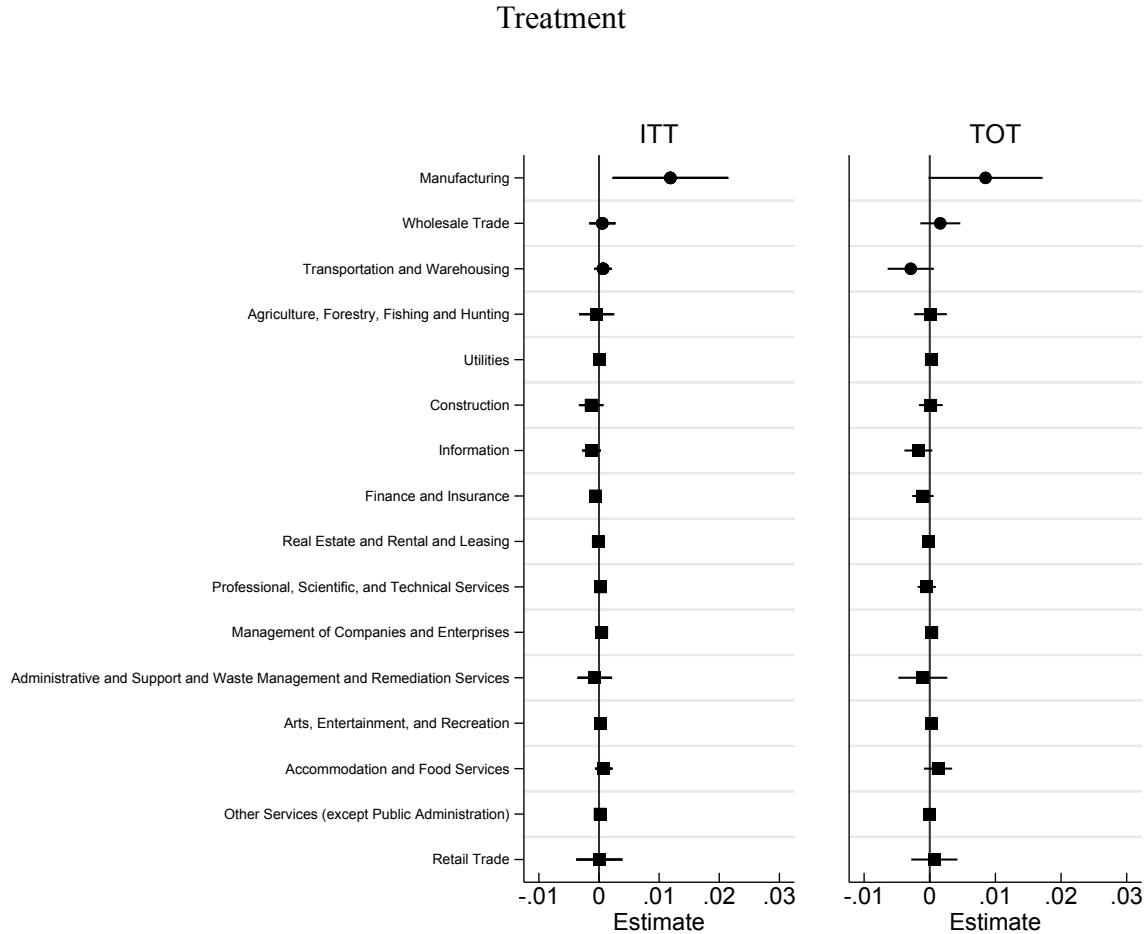
Treatment

Note: Authors' calculations. Each row comes from separate estimations of equations (1) and (2) and reports the treatment effect estimates 3 years after treatment designation. Circle markers denote target industries. Square markers denote non-target industries. Health Care and Mining are excluded because of small sample sizes. For the treatment on the treated (TOT) estimates, OLS estimates are reported. The bars around each coefficient are confidence intervals at the 95% level. Standard errors are clustered by county. Additional controls are the lagged 3-year averages of the unemployment rate and real income per capita and population growth since the most recent census, which are the 3 inputs to the distress rank. ITT is intention to treat.

5.3 Estimates for the Second Program Wave

Ideally, we would leverage the expanded eligibility of Wave 2 for the largest hiring credits to quantify program impact heterogeneity as a function of initial economic distress. Unfortunately, the randomization at the eligibility threshold underpinning the RD design for Wave 1 estimates is not as supportable for Wave 2. This wave ran from 2007 to 2013, with a significant expansion of hiring credits. As shown in table 1, the number of counties eligible for the most extensive credit amounts increased from 10 to 40. However, identifying the effects of this second program wave is difficult because the impact of the first wave of the program induces differences in the evolution of labor market outcomes that change discontinuously across the distress rank cutoff of 40. We try to control for the differences caused by the first wave of the program by estimating the effects of the second wave on a sub-sample of counties that did not have variation in credit size from 2003 to 2006. Nevertheless, there are substantial differences in covariates across the 40 distress rank cutoff in this sub-sample. Appendix figure B.4 shows a lack of balance in covariates for low bandwidths around the cutoff. Such unbalance contrasts with the previous results for Wave 1 in figure B.2. For completeness, in appendix table B.6, we show local estimates of the effects of the second wave of the program. These estimates show some significant unemployment reductions, though no employment increase.

5.4 Difference-in-Differences Estimates

Our estimate of the employment impact of hiring credits of around 4% contrasts with previous hiring credits studies, which have tended to find increases of 1%, and in some cases, no impact. Two possibilities for this discrepancy are methodology and setting. The program we study targets severely distressed areas where, as discussed earlier, hiring credits can theoretically be more impactful than in the average targeted location. Besides the setting, previous studies have relied on difference-in-differences techniques that have the potential to be biased (positively or negatively) relative to an RD approach. To assess the extent to which methodological differences underpin our larger impact estimates, we perform a difference-in-differences estimation on the program.

Our basic specification is

$$Y_{ct} = \beta_0 + \gamma_c + \gamma_t + \delta_c \times t + \sum_{k=0}^{K} \theta_k tier1_{t-k} + X_{c,t-K-1}\beta + \varepsilon_{ct}. \tag{3}$$

Here, Y_{ct} is the outcome of interest for county c at time t. γ_c and γ_t are county and year effects intended to capture permanent differences across counties and common shocks that affect all counties each year. The variable $tier1_{t-k}$ is a dummy variable equal to 1 whenever a county is assigned to Tier 1. The coefficients of interest, θ_k, capture the contemporaneous and lagged effects of Tier 1 status on the outcome variables.

This difference-in-differences strategy is valid only if unobservables have a similar evolution over time across tiers. We also include control variables X_{ct} to allow for some county heterogeneity. We include lags of population growth, real income per capita, and the unemployment rate that constitute the distress rankings' inputs. These controls address the possibility of mean reversion in outcomes or counties evolving heterogeneously because of different initial conditions before the beginning of the program (Heckman, Lalonde, and Smith, 1999). We also allow for county-specific linear time trends.

In Table 6, we present the difference-in-differences results with full controls three years after program assignment for comparison with our RD results from above. Log employment and employment to population show no impact, while the unemployment rate shows a reduction of around 0.5 percentage points. The complete estimates are in appendix table B.2.

The RD estimates of a positive and relatively large employment effect we present above imply that the difference-in-differences estimates of a null program effect are materially downward biased. This bias is consistent with the control function plotted in panel a of Figure 5 that is slightly upward sloping: Within the treatment or control group, employment in higher-ranked less-distressed counties tended to grow faster. In contrast, the control function for unemployment rates shows some tendency for convergence between more- and less-distressed counties independent of the program. In line with this pattern, the difference-in-differences unemployment estimates are not notably biased relative to the RD unemployment estimates.

Table 6: Difference-in-Differences Estimates 3 Years after Treatment

	Dependent Variable	
Log Employment	Employment to Population	Unemployment Rate
0.001	-0.001	-0.466**
(0.013)	(0.003)	(0.179)

Clustered standard errors are in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. Sample size is 546 observations for 42 counties. The table depicts treatment effect estimates from equation (3), 3 years after treatment designation. All rows include county effects, time effects, linear time trends interacted with county dummies, and lagged 3-year averages of the unemployment rate, real income per capita and population growth since the most recent census, which are the 3 inputs to the distress rank. Standard errors are clustered by county to allow for serial correlation in the error term ε_{ct} within counties (Bertrand, Duflo, and Mullainathan, 2004). P-values are calculated using a wild cluster bootstrap (Cameron and Miller, 2015) to account for the small number of clusters. We report clustered standard errors and significance tests from the bootstrap p-values with 500 replications. Full estimation results are in Table B.2 of the appendix.

5.5   Program Costs

While these hiring subsidies positively affect the labor market, a complete evaluation must consider the policy's cost. The cost accounting has some complications detailed below, but low wastage rates in this setting make the cost-per-job created look favorable. Lane and Jolley (2009) produced an evaluation of the incentive programs for the North Carolina General Assembly shortly after the Lee program had ended and morphed into the similarly conceived Article 3J program. In its entirety, the program generated $2.1 billion in tax credits through its 11-year run from 1996 to 2006. They estimate that 35% of these credits would never be used. Such lack of usage could result from future net job losses at incented firms, leading to clawback of previously generated credit value. Insufficient future tax liability could also be a factor, as the taken credits could not exceed 50% of a firm's annual tax liability, and unused credits could be carried forward only for five years. This paper's focus, the job creation credits, constituted 17% of the total amount of generated credits.[19] Though the program was more generous on a per-hire basis for the most distressed Tier 1 counties,

---

[19]The breakdown of credits generated over the life of the Lee program was job creation (17%), M&E investment (66%), R&D (15%), training (1%), and central administrative office (1%).

firms in those counties received only 14% of generated credit dollars. Of the credits, 25% went to Tier 2 counties, and more than half of all credit dollars went to firms in the least distressed Tier 3 counties, given their greater size and hiring rates.[20] Putting these figures together, we estimate that firms in our treated Tier 1 counties took job-creation hiring credits of $56.6 million (or $5.1 million per year) for 6,751 hires.[21]

In Tier 1, on average, about 24 firms per year generated some hiring credits for the 2001 to 2006 period when granular data were available. Several factors in the design of the program may have contributed to lower usage. First, a firm had to operate in a target industry and have at least five existing full-time employees. Second, it had to provide health insurance for full-time positions and not have recently violated environmental standards or safety requirements. Third, the new hire needed to be for a full-time job and receive pay above the county average wage. Finally, the business would need sufficiently positive tax liability, as credits could not exceed 50% of a taxpayer's total annual corporate income and franchise tax liability. A survey of North Carolina businesses found that incentives ranked as only the 12th most important factor in company location decisions, behind factors like access to skilled labor, highway access, tax rates, and regulator climate. Of surveyed executives at incented companies, 62% were unaware their company received an incentive. However, the share of surveyed firms in the least distressed tier where credits per hire were only $500 was not reported (Lane and Jolley, 2009).

The issue of wastage, the share of expenditures for hires that would have taken place without the policy change, is also complicated in the specific program we study. The tax expenditures flowing to the most distressed counties were relatively small, given the program impacts we find. Overall, tax expenditures under the program were much larger, as most benefits went to firms in economically robust counties where similar hiring and investment would likely have occurred without the program.

---

[20]This breakdown is based on program years 2002 to 2006.

[21]We arrive at these estimates for Tier 1 as follows: We have only aggregated figures for 1996 to 2000, so we multiply the total credit-generated figure for 1996 to 2000—$1.064 billion—by the 0.17 share for job-creation hiring credits. Then we multiply by 0.14 for the Tier 1 share and 0.65 for generated credits not taken. We add to this amount the values from more granular data on credits generated for 2001 to 2006 multiplied by 0.65 for generated credits not taken and divide by 11 for the per-year figure.

To calculate a cost-per-job, we first generate the program impact on employment levels using the OLS log employment TOT specification from Table 3. We obtain the predicted increase in log employment for a county assigned to treatment three years before but not in the intervening years (a 0.9% increase). Because the expenditure data are tier-wide, but the program impact is estimated only on non-defier counties, we need to assume a program benefit for that population. Because the defier Tier 1 counties tend to be less distressed, the program impact is likely smaller in those counties. We estimate that the differential generosity in credit size for Tier 1 relative to Tier 2 increased employment levels by 5,223–5,823 jobs three years later, assuming an impact in defier counties ranging from zero employment increase to an employment increase equal to that in non-defier counties. Adjusting the Tier 1 hiring credit expenditure data down to account for the total credit size ($12,500) versus the differential credit generosity ($9,000) yields a cost per job of around $7,400 in nominal dollars (around $12,000 in 2020 dollars). This figure contrasts with a study by Luger and Bae (2005) of the Lee Act for the 1999 program year, where they simulated reductions in the user cost of labor from the hiring credits to project induced employment effects. They project a cost per job in Tier 1 of $14,000 though they note that this estimate is highly sensitive to assumptions about the average wage. [22] According to Bartik (2020), the benefit of a created job ranges from $12,000 to $48,000, making these hiring subsidies cost-effective in isolation, though potentially more costly than alternative job-creation strategies like infrastructure investment.

Ideally, we would estimate wastage from this program by comparing our employment-level estimate of around 5,500 jobs created with the 6,751 hiring credits claimed in Tier 1 throughout the program. This comparison would imply an atypically low windfall wastage rate of approximately 20%. The contrast is unclear, though, for two reasons pushing in opposite directions. First, the program's impact theoretically combines direct and spillover job creation. That said, Table 5 and Figure 6 indicate that spillover job creation was minimal. Second, because our program impact is for the differential credit generosity between Tiers 1 and 2 rather than between a Tier 1 credit

---

[22]Specifically, Luger and Bae (2005) assume a $30,000 wage. Their simulated increase in labor demand in Tier 1 counties more than doubles under the assumption of a $20,000 average wage, closer to the prevailing Tier 1 county wages at the time, implying significantly lower projected cost per induced job.

and nothing, the *total* impact of Tier 1 designation is ostensibly larger than the differential impact. Though we cannot arrive at a precise figure, it seems that most of the claimed credits were for hires induced by the program. Such low wastage contrasts with more broad-based employment subsidy programs where Bartik (2001) puts typical windfall wastage rates at around 70% and sometimes as high as 90%, drastically inflating the effective costs per job created. Moreover, if little of the credits claimed by firms in the more economically robust higher tiers induced job creation, the overall wastage rates for the hiring subsidies across the entire state would be closer to those cited figures. For comparison, Slattery (2020) finds that discretionary subsidies—those that target a specific firm—average around $100,000 per job promised.

5.6   Validity

While using an RD strategy should provide confidence in our estimates' internal validity, some bias is still possible. Because subsidies are place-of-work based, spillover effects from any cross-county commuting induced by the program are possible. In instances where Tier 2 counties border Tier 1 counties, cross-county commuting would entail overestimating employment impacts and underestimating unemployment impacts.

Some aspects of the program under study may limit the results' external validity. Most important is the targeting of the credits to areas experiencing long-run economic distress. We would not expect the same size credit to have as large an impact on the average area as we find in distressed areas. Even in distressed areas, we cannot credibly project the effect of a much smaller per-job credit, as the benefit per dollar may be non-linear. Further, we do not estimate the impact of some versus no credits but, rather, a big versus a small credit. However, under the assumption of a decreasing marginal product of labor, we would expect the former comparison to yield impacts at least as large as what we find.

Finally, pairing the hiring credits with other incentives for investment and R&D, while not atypical, means that our estimates are full program effects rather than pure hiring credit impacts. Because the discontinuities are absent or less pronounced across tiers and take-up of the other

incentives is low in the counties on which we focus, we believe we are isolating a hiring credit effect.[23]

The persistence of hiring credit impact on the labor market is another dimension to consider. In Table B.7 in the appendix, we re-estimate equation (1) to look at program impacts out to six years subsequent to initial eligibility. Increased employment is still evident, even six years later, and of similar magnitude to the shorter-run three-years-later estimate.[24] The impact on employment-to-population ratios is more precisely estimated and confirms that employment effects do not diminish and perhaps even increase in the medium run, with point estimates rising by about 50% between three and six years later. Relative improvements in unemployment rates in treated counties demonstrate less persistence than employment increases appearing to wane beyond the three-year mark.

6   Discussion

Hiring tax credits are a popular tool implemented in various US states to lure businesses or revitalize moribund local economies. Assessing their efficacy is challenging, though, as their implementation is typically expressly endogenous to local conditions and expected prospects. We use the unusual institutional features of a program in the state of North Carolina to get causal estimates of the impact of hiring tax credits on employment and unemployment rates. Our RD ITT estimates of the impact on employment show a substantial boost from the program of 3.6-4.5%. In contrast, difference-in-differences estimates show no effect, in line with previous studies of hiring credits using cross-state comparisons. The downward bias in these latter estimates, relative to the more credible RD ones, highlights the importance of accounting for time-varying unobservables. These

---

[23]Under the Lee program, there was a 5% credit for R&D expenditures which was the same across all tiers. Over 90% of these credits went to firms in Tier 3 counties. There was also a 7% credit for machinery and equipment investment which was the same across all tiers. While any size investment qualified in Tier 1, only investments over $100 thousand in Tier 2 received the credit. Over three-quarters of the M&E credits went to firms in Tier 3 counties (Fain, 2005).

[24]Relative to the estimations presented in Table 2, the window around each treatment designation is expanded forward by three years, increasing the sample size. The expanded window means county fixed-effect coefficients change, influencing the coefficients capturing the one to three years later impact estimates. That is why the estimates in Table B.7 can and do differ slightly from what appears in Table 2.

unobservables may affect counterfactual county performance. This is apparent in the RD graphs of Figure 5. The absolute outcome for treated counties is close to no employment growth, which contrasts with substantial contraction in similarly distressed control counties.

Cahuc et al. (2019) find consistent evidence of positive employment impacts of hiring credits at eligible firms, an impact estimate most analogous to our target industry results. Their estimates range from 0.8% to 2.3% for a credit of about one-fourth the size of what we study. Assuming the impact is linear in credit size, that finding is consistent with our target-industry employment growth estimates of around 7%.

We also find impacts on unemployment, with treated counties—those whose firms were eligible for large hiring tax credits—experiencing about 0.4 percentage points lower unemployment rates than under a counterfactual program offering much smaller credits. The program's windfall wastage rate appears to have been low, possibly due to the targeting of distressed areas where hiring in the program's absence would have been limited. So even though inducing more economic activity in low-productivity distressed areas may require large incentives, the bulk of incentive expenditures are changing firms' behavior rather than being windfall payments. While hiring subsidies may not reverse regional divergence, they have a significant positive impact. They can be designed cost-effectively if eligibility is restricted to the most distressed labor markets.

References

Amior, Michael and Alan Manning (2018), "The Persistence of Local Joblessness." *American Economic Review*, 108, 1942–70.

Bartik, Timothy J. (2001), *Jobs for the Poor: Can Labor Demand Policies Help?* W.E. Upjohn Institute for Employment Research.

Bartik, Timothy J. (2020), "Smart Place-Based Policies Can Improve Local Labor Markets." *Journal of Policy Analysis and Management*, 39, 844–851.

Behaghel, Luc, Adrien Lorenceau, and Simon Quantin (2015), "Replacing churches and mason lodges? tax exemptions and rural development." *Journal of Public Economics*, 125, 1–15.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119, 249–275.

Busso, Matias, Jesse Gregory, and Patrick Kline (2013), "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review*, 103, 897–947.

Cahuc, Pierre, Stéphane Carcillo, and Thomas Le Barbanchon (2019), "The Effectiveness of Hiring Credits." *Review of Economic Studies*, 86, 593–626.

Cahuc, Pierre, Stéphane Carcillo, and André Zylberberg (2014), *Labor Economics*. The MIT Press.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik (2015), "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association*, 110, 1753–1769.

Cameron, Collin A. and Douglas L. Miller (2015), "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, 50, 317–372.

Cattaneo, Matias D, Brigham Frandsen, and Rocıo Titiunik (2015), "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the US Senate." *Journal of Causal Inference*, 3, 1–24.

Cattaneo, Matias D., Rocio Titiunik, and Gonzalo Vazquez-Bare (2016), "Inference in Regression Discontinuity Designs Under Local Randomization." *Stata Journal*, 16, 331–367.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010), "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *The Quarterly Journal of Economics*, 125, 215–261.

Cerqua, Augusto and Guido Pellegrini (2014), "Do Subsidies to Private Capital Boost Firms' Growth? A Multiple Regression Discontinuity Design Approach." *Journal of Public Economics*, 109, 114–126.

Chay, Kenneth Y., Patrick J. McEwan, and Miguel Urquiola (2005), "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools." *The American Economic Review*, 95, 1237–1258.

Chirinko, Robert S. and Daniel J. Wilson (2016), "Job Creation Tax Credits, Fiscal Foresight, and Job Growth: Evidence from U.S. States." CESifo Working Paper Series 5771, CESifo Group Munich.

Coate, Patrick and Kyle Mangum (2019), "Fast Locations and Slowing Labor Mobility." Working Papers 19-49, Federal Reserve Bank of Philadelphia.

Dong, Yingying and Arthur Lewbel (2015), "Identifying the Effect of Changing the Policy Threshold in Regression Discontinuity Models." *The Review of Economics and Statistics*, 97, 1081–1092.

Fain, James T. (2005), "William S. Lee Act: 2005 Assessment of Results." Technical Report dp1357, State of North Carolina Department of Commerce: Division of Policy, Research & Strategic Planning.

Fain, Jim (2001), "William S. Lee Quality Jobs and Business Expansion Act: Assessment of Results." Technical report, North Carolina Department of Commerce: Division of Economic Policy & Research.

Faulk, Dagney (2002), "Do State Economic Development Incentives Create Jobs? An Analysis of State Employment Tax Credits." *National Tax Journal*, 55, 263–280.

Freedman, Matthew (2013), "Targeted Business Incentives and Local Labor Markets." *Journal of Human Resources*, 48, 311–344.

Gelman, Andrew and Guido Imbens (2019), "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics*, 37, 447–456.

Gobillon, Laurent, Thierry Magnac, and Harris Selod (2012), "Do Unemployed Workers Benefit from Enterprise Zones? The French Experience." *Journal of Public Economics*, 96, 881–892.

Grijalva, Diego and David Neumark (2014), "State Hiring Credits and Recent Job Growth." *FRBSF Economic Letter*, 05.

Heckman, James J., Robert J. Lalonde, and Jeffrey A. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics* (O. Ashenfelter and D. Card, eds.), volume 3, chapter 31, 1865–2097, Elsevier.

Huttunen, Kristiina, Jukka Pirttilä, and Roope Uusitalo (2013), "The Employment Effects of Low-Wage Subsidies." *Journal of Public Economics*, 97, 49–60.

Jolley, G Jason, E Brent Lane, and Sharon R Paynter (2013), "Who Benefits from Job Creation Tax Credits?" *Coastal Business Journal*, 12, 135–145.

Kline, Patrick and Enrico Moretti (2013), "Place Based Policies with Unemployment." *American Economic Review*, 103, 238–43.

Kline, Patrick and Enrico Moretti (2014), "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics*, 6, 629–662.

Ku, Hyejin, Uta Schönberg, and Ragnhild C. Schreiner (2020), "Do Place-Based Tax Incentives Create Jobs?" *Journal of Public Economics*, 191, 104105.

Lane, Brent and G. Jason Jolley (2009), "An Evaluation of North Carolina's Economic Development Incentive Programs: Final Report." Technical report, Chapel Hill: University of North Carolina at Chapel Hill.

Lee, David S. and Thomas Lemieux (2010), "Regression discontinuity designs in economics." *Journal of Economic Literature*, 48, 281–355.

Luger, Michael I and Suho Bae (2005), "The Effectiveness of State Business Tax Incentive Programs: The Case of North Carolina." *Economic Development Quarterly*, 19, 327–345.

Neumark, David (2013), "Spurring Job Creation in Response to Severe Recessions: Reconsidering Hiring Credits." *Journal of Policy Analysis and Management*, 32, 142–171.

Neumark, David (2020), "Place-Based Policies: Can We Do Better Than Enterprise Zones?" *Journal of Policy Analysis and Management*, 39, 836–844.

Neumark, David and Diego Grijalva (2017), "The Employment Effects of State Hiring Credits." *ILR Review*, 70, 1111–1145.

Neumark, David and Jed Kolko (2010), "Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program." *Journal of Urban Economics*, 68, 1–19.

Neumark, David and Timothy Young (2019), "Enterprise Zones, Poverty, and Labor Market Outcomes: Resolving Conflicting Evidence." *Regional Science and Urban Economics*, 78.

Papke, Leslie E. (1994), "Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program." *Journal of Public Economics*, 54, 37–49.

Program Evaluation Division (2015), "Final report to the joint legislative program evaluation oversight committee." Technical Report 2015-11, North Carolina General Assembly.

Sestito, Paolo and Eliana Viviano (2018), "Firing Costs and Firm Hiring: Evidence from an Italian Reform." *Economic Policy*, 33, 101–130.

Slattery, Cailin (2020), "Bidding for Firms: Subsidy Competition in the U.S." *Working Paper*.

Wong, Vivian C, Peter M Steiner, and Thomas D Cook (2013), "Analyzing Regression-Discontinuity Designs With Multiple Assignment Variables: A Comparative Study of Four Estimation Methods." *Journal of Educational and Behavioral Statistics*, 38, 107–141.

Zabek, Mike (2019), "Local Ties in Spatial Equilibrium." *Finance and Economics Discussion Series. Washington: Board of Governors of the Federal Reserve System*, 2019-080.

Online Appendix - Not For Publication

A   Additional program details

The first iteration of North Carolina's non-discretionary tax incentive program, denoted as wave 0 in Table 1, began in 1988. The state Department of Commerce was tasked with ranking counties each year from 1 to 100 based on economic distress, which legislation defined as the combination of a high unemployment rate and low per capita incomes. Specifically, counties were ranked separately by unemployment rate and income per capita and then the two rankings were summed to produce an overall distress rank (potentially including ties). Businesses in the 20 most distressed counties were eligible for a $2,800 tax credit for each new full time employee hired. The number of eligible counties was progressively increased, and had reached 50 by the time the program ended in 1995. We were unable to find data on this program's implementation beyond the identity of the 50 counties eligible for credits in its final year which we use as a control variable.

A revamped program was launched in 1996 known as Article 3A or the William S. Lee program which is the focus of our study and referred to as wave 1. It continued the use of a county ranking scheme and added population growth as an input to the ranking process. It extended tax credits to all counties which varied in size based on groups of counties known as tiers. The largest credits of $12,500 were available to the 10 most distressed counties designated as Tier 1. Firms in less distressed counties could receive credits between $500 and $4,000 per new hire. Over the course of this program, the number of counties eligible for the largest credit size was increased as low population and high poverty rates were added as overrides of the distress ranking system, with 28 counties designated Tier 1 and eligible for the largest credits by the final year of the program in 2006.

The William S. Lee program was itself replaced in 2007 by the Article 3J program referred to as wave 2. This latter program operated in similar fashion to its predecessors, but with some changes to the credit eligibility formulas. Tier 1 - the most distressed - expanded to contain 40 counties eligible for credits of $12,500. The next most distressed 40 counties were in Tier 2 and could receive $5,000

credits and the highest performing 20 counties in Tier 3 could receive $750 credits. The distress ranking formula was amended to incorporate property value per capita alongside unemployment rate, income, and population growth. In 2014, the Article 3J program ended and was not replaced (Program Evaluation Division, 2015).

Table A.1: Distress Ranking Example - 2005

County	Population growth (1)	rank (2)	Income per cap (3)	rank (4)	Unemp. rate (5)	rank (6)	Distress sum (7)	rank (8)	Pop. (9)	Poverty (10)	Override (11)	Tier (12)
Vance	-0.54%	96	$21,697	74	12.43	100	270	1	44,134	20.50	0	1
Halifax	-0.27%	86	$20,132	91	8.83	90	267	2	56,533	23.90	0	1
Scotland	-0.26%	85	$21,083	82	10.96	99	266	3	35,089	20.56	0	1
Hyde	-2.14%	100	$19,694	93	7.16	70	263	4	5,854	15.44	0	1
Washington	-0.85%	98	$20,926	85	7.69	78	261	5	13,507	21.76	0	1
Edgecombe	-1.39%	99	$22,373	64	9.59	96	259	6	55,583	19.59	0	1
Richmond	-0.21%	84	$21,195	81	9.37	93	258	7	46,053	19.56	0	1
Martin	-0.71%	97	$21,520	77	8.09	83	257	8	24,940	20.19	0	1
Warren	0.36%	62	$17,947	99	9.38	94	255	9	20,252	19.45	0	1
Yancey	-0.04%	80	$19,621	94	7.86	80	254	10	17,546	15.76	0	1
Mitchell	-0.09%	82	$20,004	92	7.59	77	251	11	15,770	13.83	0	2
Bertie	0.13%	76	$20,695	87	8.50	88	251	11	20,013	23.46	1	1
Anson	-0.35%	91	$22,536	60	10.25	97	248	13	25,690	17.77	0	2
Caswell	0.08%	78	$21,537	76	8.37	86	240	14	23,738	14.40	0	2
Robeson	0.65%	53	$18,238	98	8.31	85	236	15	125,394	22.81	0	2
North Carolina	1.33%		$27,644		5.47				8,418,493	12.23		

Note: Example of the computation of the distress ranking and tier assignment for the 15 most distressed in 2005 out of the 100 counties in North Carolina. The three inputs to the rankings are recent county population growth, income per capita, and unemployment rates. These are ranked separately from 100 to 1 in columns (2), (4) and (6) and then summed and ordered for the overall distress rank in column (8). The ten lowest ranked/most distressed counties are assigned to Tier 1. Counties ranked 11 or higher start being assigned to Tier 2, unless they trip an override based on a combination of low population and a high poverty rate that re-designates them as Tier 1. Overall values for North Carolina's 100 counties at the bottom.

B Additional tables and figures

Figure B.1: Relationship between the Distress Rank and Covariates Before the Beginning of the

Program

(a)

(b)
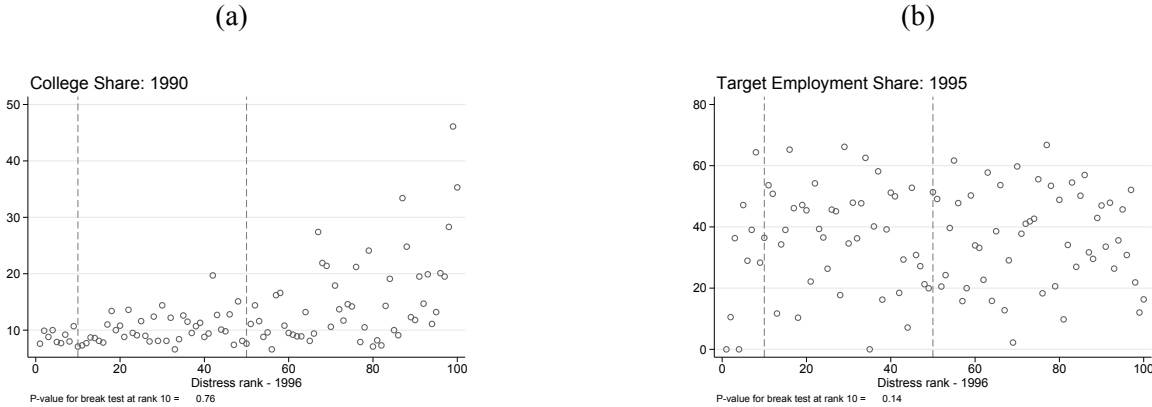
Note: College share data are from the Census. Employment data are from the Bureau of Labor Statistics. County-level economic indicators are arrayed by initial distress rank at the outset of the first wave of the program. Vertical lines denote the thresholds where credit size jumps discontinuously. The reported p-value comes from a linear regression of each economic indicator on the distress rank and a dummy for a break at 10, restricted to the 50 most distressed counties.

Table B.1: Treatment Effect Derivative (TED) Estimates 3 Years After Treatment

Dependent Variable	Log Employment		Employment to Population		Unemployment Rate	
Treatment effect	0.030*	0.036**	0.010*	0.012**	-0.75**	-0.51**
	(0.017)	(0.017)	(0.0052)	(0.0049)	(0.32)	(0.23)
Control function	0.0020***	0.0016**	0.00040***	0.00024	0.0020	-0.019**
	(0.00066)	(0.00074)	(0.00015)	(0.00016)	(0.0076)	(0.0091)
Treatment effect	0.042	0.044	0.013	0.013	-0.63	-0.61
	(0.031)	(0.031)	(0.0087)	(0.0080)	(0.50)	(0.37)
Control function	0.0019***	0.0016**	0.00038**	0.00024	0.0016	-0.018**
	(0.00065)	(0.00073)	(0.00014)	(0.00015)	(0.0074)	(0.0082)
TED	0.0034	0.0025	0.00068	0.00018	0.031	-0.035
	(0.0061)	(0.0066)	(0.0015)	(0.0016)	(0.075)	(0.081)
Covariates	No	Yes	No	Yes	No	Yes

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. N=2,779. The top panel repeats the 3 years later estimates from table 2 and shows the estimated linear control function in the distress rank. The bottom panel adds the interaction between treatment status and the running variable to the specification, allowing the linear control function to change slope at the threshold. The coefficient on the interaction term is the treatment effect derivative (TED) discussed in Dong and Lewbel (2015). Standard errors are clustered by county. All rows include fixed effects for year, county, time since tier designation, and prior treatment history. Columns with covariates add as additional controls the lagged 3-year averages of the unemployment rate and real income per capita and population growth since the most recent census, which are the 3 inputs to the distress rank.

Table B.2: Difference-in-Differences Estimates

Dependent Variable	Log Employment		Employment to Population		Unemployment	
	(1)	(2)	(3)	(4)	(5)	(6)
Tier 1	-0.007	0.001	-0.001	0.001	0.072	0.077
	(0.009)	(0.008)	(0.002)	(0.002)	(0.165)	(0.160)
Lag Tier 1	-0.021**	-0.016*	-0.004*	-0.004*	-0.121	-0.051
	(0.009)	(0.009)	(0.002)	(0.002)	(0.169)	(0.186)
Lag 2 Tier 1	-0.014	-0.011	-0.003	-0.002	-0.293**	-0.246**
	(0.009)	(0.008)	(0.002)	(0.002)	(0.117)	(0.112)
Lag 3 Tier 1	-0.002	0.001	-0.001	-0.001	-0.518***	-0.466**
	(0.014)	(0.013)	(0.003)	(0.003)	(0.189)	(0.179)
Lag 4 Population growth		-0.000		-0.001		0.042
		(0.003)		(0.001)		(0.030)
Lag 4 Real Income per capita		0.020		-0.003		0.655
		(0.045)		(0.012)		(0.694)
Lag 4 Unemployment Rate		-0.006**		-0.001		-0.111**
		(0.003)		(0.001)		(0.046)
R^2	0.998	0.998	0.981	0.982	0.810	0.826
N	588	546	588	546	588	546
$Counties$	42	42	42	42	42	42
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County trends	Yes	Yes	Yes	Yes	Yes	Yes

Clustered standard errors are in parentheses

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Note: Authors' calculations. Difference-in-differences estimates of equation (3). Standard errors are clustered by county. P-values for significance tests are calculated using a wild cluster bootstrap with 500 replications to account for the small number of counties. All columns include county and time effects, and linear time trends interacted with county dummies.

Table B.3: Regression Discontinuity TOT IV Estimates - Main Outcomes

Dependent Variable	1 year later	2 years later	3 years later
Log Employment	-0.065*	0.038	0.072**
	(0.038)	(0.031)	(0.031)
Employment to Population	-0.016	0.010	0.023**
	(0.011)	(0.008)	(0.009)
Unemployment Rate	-0.154	-1.049**	-1.144*
	(0.633)	(0.498)	(0.609)

Clustered standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. N=770. Each row comes from a separate estimation of equation (2) and reports the treatment effect estimates θ_k at 1, 2, and 3 years after treatment designation. Standard errors are clustered by county. All rows include lags of the unemployment rate, income per capita, and population growth as controls. Instrumental variable (IV) estimates corresponding to a fuzzy RD design, which labels all Tier 1 counties as being treated and instruments for treatment with the distressed rank assignment rule. TOT is Treatment on the Treated.

Figure B.2: Local Estimates Window Selection
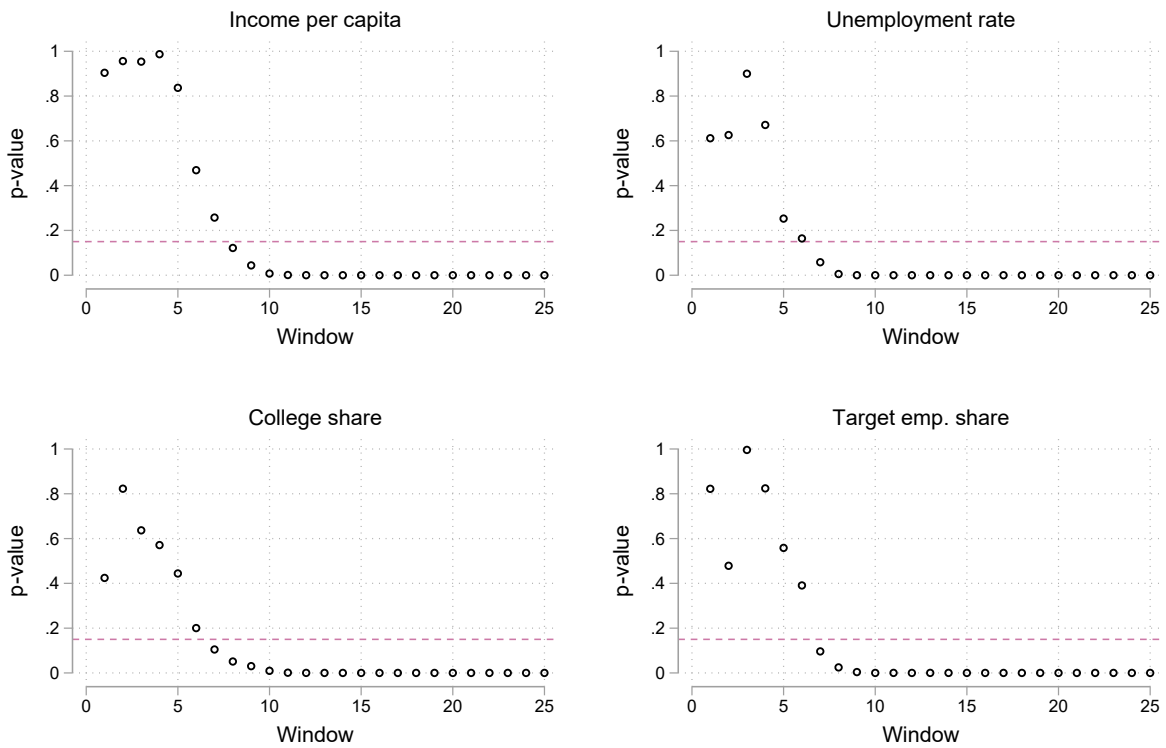
Note: Authors' calculations. Covariate balance tests by window size against the 15% level threshold recommended by Cattaneo et al. (2015).

Figure B.3: Stability of 3-Years-Later Local Estimates for Different Window Widths
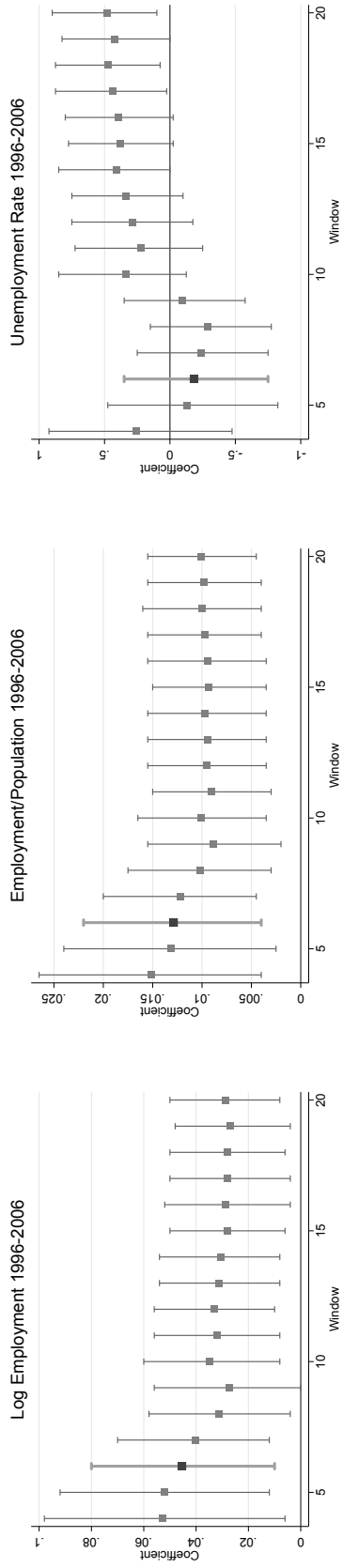
Note: Authors' calculations. Difference in mean outcomes for treated and control counties, for different bandwidths of distress ranks around the policy threshold. For unemployment, the estimates are adjusted for linear dependence in the distress rank. The bars are 95% confidence intervals obtained by randomization inference with a 1000 repetitions. The preferred estimates, for a six rank bandwidth, are displayed with thicker lines and markers. The bandwidth selection diagnostics are in figure B.2.

Table B.4: Local Estimates of Effect 3 Years after Treatment, Excluding Northampton County in

2005

Time Range	Window	Dependent Variable						N
		Log Employment		Employment Population		Unemployment Rate		
1996-2006	6 ranks	0.038**	[0.024]	0.012**	[0.016]	-0.230	[0.460]	93
1996-2006	10 ranks	0.031**	[0.013]	0.009***	[0.004]	-0.536**	[0.031]	159
1996-2006	20 ranks	0.025**	[0.027]	0.009***	[0.001]	-0.731***	[0.000]	239

P-values from randomization inference with 1000 replications in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. Difference in mean outcomes for treated and control counties, for different bandwidths of distress ranks around the policy threshold, three years after treatment. Estimates exclude Northampton county in 2005 which is a visual outlier in Figure 5.

Figure B.4: Local Estimates Window Selection. Second Program Wave.
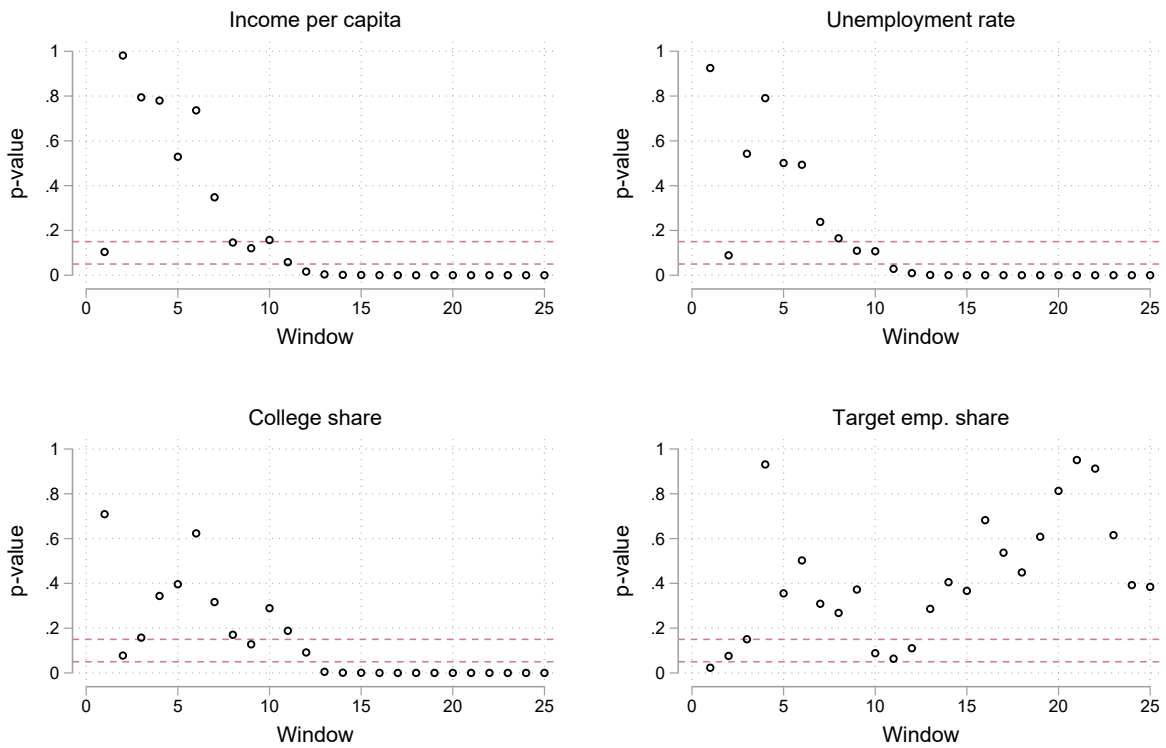
Note: Authors' calculations. Covariate balance tests by window size against the 15% level threshold recommended by Cattaneo et al. (2015).

Table B.5: Local Estimates of Effect 3 Years after Treatment Allowing for Dependence on the

Ranking Variable

Years Ahead	Polynomial Order	Dependent Variable			N
		Log Employment	Employment Population	Unemployment Rate	
1	0	0.002	0.001	-0.279	118
		[0.861]	[0.682]	[0.122]	
		{−386.17}	{−701.37}	{321.05}	
1	1	0.013	0.003	-0.416**	118
		[0.121]	[0.253]	[0.017]	
		{−387.88}	{−700.60}	{324.46}	
1	2	0.034***	0.008***	-0.035	118
		[0.000]	[0.002]	[0.846]	
		{−386.97}	{−698.57}	{327.51}	
2	0	0.022	0.006	-0.271	106
		[0.145]	[0.145]	[0.308]	
		{−260.20}	{−534.81}	{359.22}	
2	1	0.034**	0.008**	-0.129	106
		[0.010]	[0.047]	[0.583]	
		{−258.66}	{−531.66}	{358.05}	
2	2	0.063***	0.018***	-0.062	106
		[0.000]	[0.000]	[0.810]	
		{−255.51}	{−528.70}	{362.01}	
3	0	0.045**	0.013***	-0.237	94
		[0.010]	[0.008]	[0.415]	
		{−191.45}	{−440.79}	{333.53}	
3	1	0.080***	0.019***	-0.182	94
		[0.000]	[0.000]	[0.520]	
		{−189.00}	{−437.46}	{333.11}	
3	2	0.105***	0.027***	0.247	94
		[0.000]	[0.000]	[0.388]	
		{−185.39}	{−434.42}	{336.45}	

Coefficient p-values from randomization inference with 1000 replications in brackets.

Akaike information criteria values in braces.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. Differences in outcomes for treated and control counties, after adjusting for a polynomial on the distress rank on each side of the cutoff.

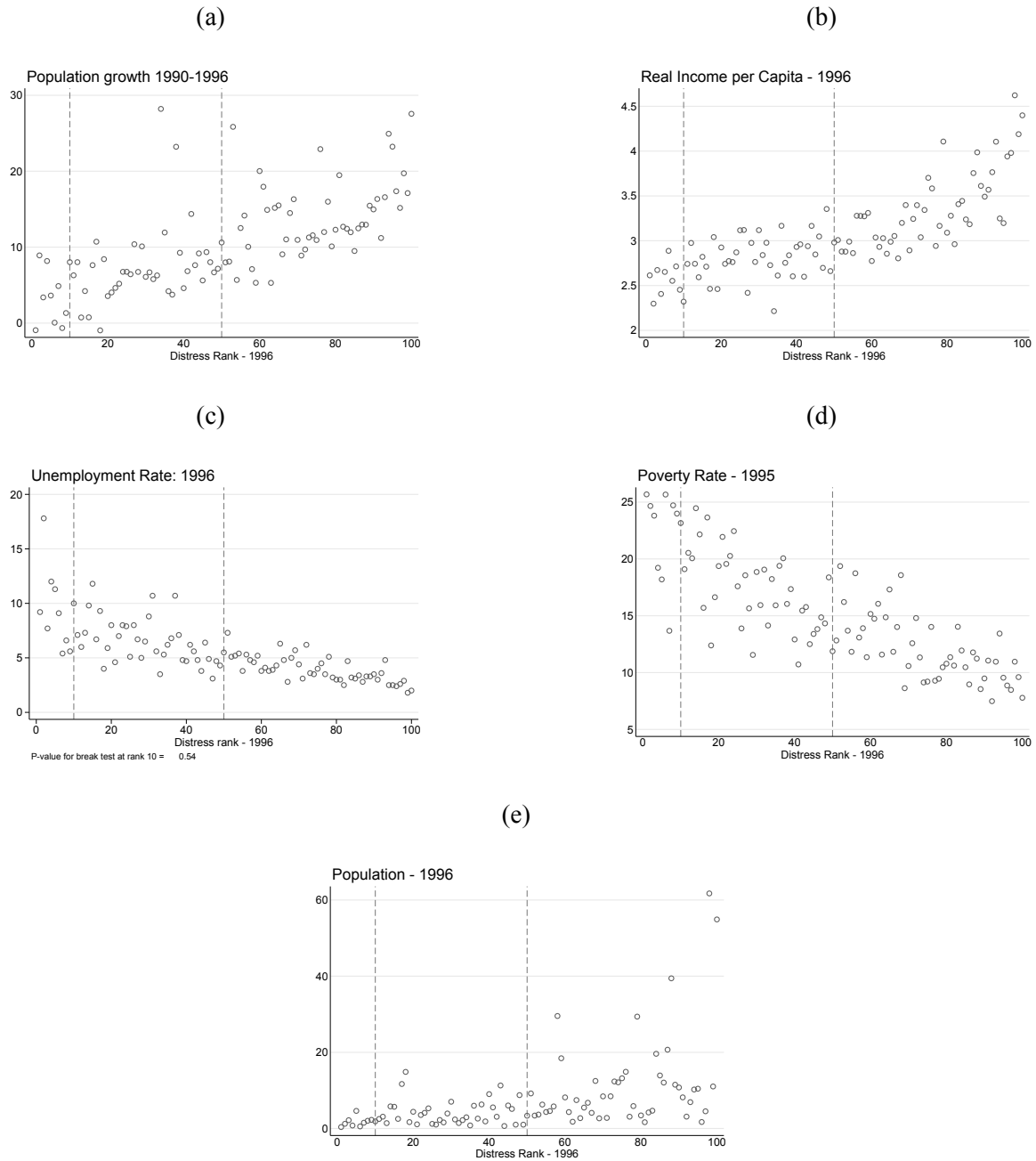Table B.6: Local Estimates, Second Program Wave: 1,2 and 3 Years after Treatment

Window	Years	Dependent Variable			N
		Log Employment	Employment to Population	Unemployment Rate	
6 ranks	1 Year	0.005	0.001	-0.409*	42
		[0.628]	[0.644]	[0.076]	
	2 Years	0.000	0.000	-0.602*	42
		[0.967]	[0.925]	[0.066]	
	3 Years	0.005	0.001	-0.712**	42
		[0.750]	[0.824]	[0.035]	
10 ranks	1 Year	-0.008	-0.002	-0.241	73
		[0.272]	[0.325]	[0.122]	
	2 Years	-0.005	-0.000	-0.451*	73
		[0.689]	[0.912]	[0.057]	
	3 Years	-0.005	-0.000	-0.671***	73
		[0.739]	[0.945]	[0.006]	
20 ranks	1 Year	0.003	0.001	-0.315***	156
		[0.641]	[0.562]	[0.002]	
	2 Years	0.003	0.002	-0.371**	156
		[0.727]	[0.481]	[0.026]	
	3 Years	0.005	0.003	-0.590***	156
		[0.588]	[0.267]	[0.001]	

P-values from randomization inference with 1000 replications in brackets.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Authors' calculations. Difference in mean outcomes for treated and control counties, for different bandwidths of distress ranks around the policy threshold.

Figure B.5: Relationship between the Distress Rank and Tier designation Inputs

(a)

Population growth 1990-1996

(b)

Real Income per Capita - 1996

(c)

Unemployment Rate: 1996

P-value for break test at rank 10 =   0.54

(d)

Poverty Rate - 1995

(e)

Population - 1996

Note: Authors' calculations. County level economic indicators are arrayed by initial distress rank at the outset of the first wave of the program. Vertical lines denote the thresholds where credit size jumps discontinuously.

Table B.7: Regression Discontinuity ITT Estimates - Effect Persistence

Dependent Variable	1 year later	2 years later	3 years later	4 years later	5 years later	6 years later
Log Employment	-0.002	0.006	0.023	0.030*	0.039	0.035
	(0.012)	(0.014)	(0.015)	(0.017)	(0.024)	(0.028)
with controls	-0.002	0.007	0.028*	0.036**	0.046**	0.040
	(0.012)	(0.014)	(0.015)	(0.017)	(0.022)	(0.027)
Employment/Population	-0.001	0.002	0.008*	0.011**	0.014**	0.013*
	(0.003)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
with controls	0.000	0.003	0.010**	0.013***	0.018***	0.016**
	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Unemployment Rate	0.151	-0.354	-0.616**	-0.633*	-0.641	-0.638
	(0.255)	(0.269)	(0.281)	(0.336)	(0.443)	(0.458)
with controls	0.257	-0.235	-0.414**	-0.290	-0.204	-0.194
	(0.302)	(0.245)	(0.204)	(0.208)	(0.289)	(0.311)

Clustered standard errors in parentheses

$* p < 0.1$, $** p < 0.05$, $*** p < 0.01$

Note: Authors' calculations. N=3,541. Each row comes from a separate estimation of equation (1) and reports the treatment effect estimates θ_k at 1 through 6 years after treatment designation. All rows include fixed effects for year, county, time since tier designation, and prior treatment history, as well as a linear control function in the distress rank. Additional controls are the lagged three-year averages of the unemployment rate and real income per capita and population growth since the most recent census, which are the 3 inputs to the distress rank.