Unemployment Insurance Generosity and Aggregate Employment*

Christopher Boone[†], Arindrajit Dube[‡], Lucas Goodman[§] and Ethan Kaplan[¶]

December 20, 2016

This paper examines the impact of unemployment insurance (UI) on aggregate employment by exploiting cross-state variation in the maximum benefit duration during the Great Recession. Comparing adjacent counties located in neighboring states, we find no statistically significant impact of increasing UI generosity on aggregate employment. Our point estimates are uniformly small in magnitude, and the most precise estimates rule out employment-to-population ratio reductions in excess of 0.5 percentage points from the UI extension. We show that a moderately sized fiscal multiplier can rationalize our findings with the small negative labor supply impact of UI typically found in the literature.

*We thank Gabriel Chodorow-Reich, Thomas Hegland, Ioana Marinescu, and Jesse Rothstein for helpful comments. Dube and Kaplan acknowledge financial support from the Institute for New Economic Thinking. We wish to thank Doruk Cengiz, Bryan Hardy and Yuting Huang for excellent research assistance.

†Cornell University, 530 Statler Hall, Ithaca, NY 14853; boone@cornell.edu

‡University of Massachusetts Amherst, 1030 Thompson Hall, University of Massachusetts Amherst, Amherst, MA 01003; adube@econs.umass.edu

§Department of Economics, 3114 Tydings Hall, University of Maryland at College Park, College Park, MD 20743; goodman@econ.umd.edu

¶Department of Economics, 3114 Tydings Hall, University of Maryland at College Park, College Park, MD 20743; kaplan@econ.umd.edu

1 Introduction

During the Great Recession, existing law and new acts of Congress led to the most dramatic expansion in the generosity of unemployment insurance (UI) benefits in U.S. history[1]. In most states, eligible job losers saw their maximum benefit duration rise from the usual 26 weeks to 99 weeks. Continuously from November 2009 through March 2012, the maximum benefit duration exceeded 90 weeks when averaged across states, except for a few small lapses. In comparison, during a previous spell of extended benefits in response to the 2001 recession, this average rarely exceeded 40 (Farber and Valletta (2015)).

This unprecedented UI expansion—and its variation across states in magnitude and timing—provides a unique opportunity to study the aggregate employment effects of UI benefit duration. In this paper, we examine the effect of UI duration on aggregate employment during the Great Recession using state-level expansions and contractions in UI generosity. We use county-level monthly employment data from late 2007 until the end of 2014. We provide transparent evidence on employment dynamics around sharp and durable changes in UI benefits across counties that were otherwise very similar, and provide a reconciliation of the differences in findings across existing papers.

While a large body of research has studied the effect of UI duration on the labor supply and job search behavior of individuals, the effects of the benefit extension on aggregate employment may be quite different from the micro-based estimates. Keynesian theory predicts a positive employment effect of UI provision during recessions via stimulating aggregated demand (Summers (2010); Congressional Budget Office (2012)). In contrast, search-and-matching models suggest that extensions could raise reservation wages and lead to lower vacancies and employment (Mitman and Rabinovich (2014)). Finally, if jobs are rationed, the decreased search from increased UI generosity during downturns may have only limited effects on aggregate employment due to increased labor market tightness (the "rat race" phenomenon)—implying a smaller macro effect than micro effect (Michaillat (2012); Landais et al. (2015); Lalive et al. (2015)). Unfortunately, a small set of recent empirical papers has delivered a mixed verdict on the size of the macro effect of the policy (Chodorow-Reich and Karabarbounis (2016); Coglianese (2015); Hagedorn et al. (2015); Hagedorn et al. (2016); Johnston and Mas (2015)).

We begin by showing that the structure of UI extensions that occurred during the Great Recession makes our task quite difficult: federal policy expanded a state's UI duration automatically when unemployment in that state was high, leading to reverse causality. To address this mechanical endogeneity, we compare neighboring counties located on opposite sides of state boundaries.[2] We show that this border-county-pair

[1] The second largest increase provided a temporary increase in unemployment duration of 65 weeks in 1975 following the passage of the Special Unemployment Insurance Extension Act.

[2] This border-county-pair strategy was first used in Dube et al. (2010) to study minimum wage policies, which change discontinuously at state borders. Note that the same problem of mechanical endogeneity does not arise when studying the

(hereafter BCP) strategy substantially reduces the endogeneity problem, mitigating negative pre-existing employment trends in counties that subsequently experienced greater expansions in maximum benefit duration. In addition to OLS specifications that make use of all variation in state-level UI duration over the entire period, we also provide an instrumental variables estimate using variation induced solely by national level policy changes—namely the November 2008 expansion and the December 2013 expiration of the Emergency Unemployment Compensation (EUC) program. These national level policy changes are less endogenous to employment changes between neighboring counties than variation resulting from the movements in state-level unemployment rates. At the same time, the bite of the policy differed across state borders, which allows us to use the BCP strategy in conjunction with the IV approach. We show changes in aggregate employment during the 12 months before and after these expansion and expiration events; we also combine the data for both events to produce a pooled IV estimate.

Our main results are as follows. We find no evidence that UI benefit extensions substantially affected county-level employment. For the full sample OLS regressions, our point estimates for the effect of expanding maximum benefit duration from 26 to 99 weeks range from 0.21 to 0.43 percentage points of the employment-to-population (EPOP) ratio. These estimates are not significantly different than zero, and they allow us to rule out negative effects on EPOP greater than -0.48 percentage points at the 95% confidence level. For comparison, the total change in EPOP over the course of the Great Recession was about -3 percentage points in our sample.

Our IV estimates that specifically use variation from the national level policy changes in 2008 and 2014 reach a similar conclusion. For the 2008 IV estimation, the point estimates also indicate positive impacts on EPOP as a result of the UI expansion, but the standard errors are much larger. For the 2014 IV analysis, however, the impacts are estimated with more precision: the point estimates are -0.02 and -0.18 percentage points of EPOP, suggesting a very small negative impact on employment. When pooled over both events, our point estimates for the effect of increasing the maximum benefit duration from 26 to 99 weeks range between -0.07 and 0.14. While the IV estimates are somewhat less precise (especially for the 2008 expansion event), the most precise pooled estimate rules out effects more negative than -1.31 percentage points of EPOP from a 73-week increase in maximum benefit duration, at the 95 percent confidence level. Similarly, the estimates from the 2014 expiration event rules out effects more negative than -1.20 from the same policy change.

These conclusions are reinforced when evaluating dynamic evidence from our distributed lag specifications. For the full sample, we find that employment remained essentially unchanged over a 36 month window that includes 24 months after treatment. In particular, we see no trends prior to treatment, indicating that

effects of the minimum wage, as statutory wage rates are not directly tied to measures of state level unemployment. However, minimum wage policies can also be subject to endogeneity bias through political economic channels, and more generally may be correlated with spatially varying confounders.

neither endogeneity nor policy anticipation confound our estimates. Event studies for the 2008 introduction and 2014 expiration also show qualitatively similar results. Taking into account the micro-econometric estimates of labor supply from other studies, we back out ranges of potential Keynesian multipliers that would be consistent with our macroeconomic estimates. Our macro employment estimates are consistent with a range of positive fiscal multipliers centered near 1 when we consider typical labor supply estimates from the UI benefit expansion—as found in many of the studies using data from the Great Recession.

A number of recent papers have exploited the panel variation across U.S. states over time in benefit duration during the Great Recession to study (micro-level) labor supply behavior. Rothstein (2011) uses data from the Current Population Survey (CPS) and variation from the uneven roll-out of extended benefits across states and finds that UI extensions were responsible for an increase in unemployment of 0.2 percentage points.[3] In concurrent work using similar variation, Farber and Valletta (2015) find that the availability of extended benefits increased the unemployment rate by 0.4 percentage points. Farber et al. (2015) find similar results when they exploit variation in UI generosity that arises due to the phase-out of extended benefits in 2012-2014: the effect of UI on duration to re-employment is small. Evaluating a sudden reduction in benefits in Missouri, Johnston and Mas (2015) reach a different conclusion: they find that newly unemployed workers who are eligible for 16 fewer weeks of UI (due to starting their claim shortly after a policy change) were 10 percentage points more likely to be employed starting in the quarter immediately after the policy change took place.

In contrast to the large empirical literature on the micro-level labor supply elasticity, there are relatively fewer papers that have estimated the macro-level impact of unemployment insurance on overall employment. The papers most closely related to ours are Hagedorn, Karahan, Manovskii and Mitman (2015)—hereafter HKMM—and Hagedorn, Manovskii and Mitman (2016)—hereafter HMM. Like us, these papers use a BCP strategy; HKMM provide evidence complementary to us that the BCP strategy mitigates the endogeneity problem. However, they both estimate large negative effects of UI on aggregate employment. HKMM find that the expansion of UI during the Great Recession from 26 to 99 weeks increased the unemployment rate by 80%, which is an effect on unemployment that is roughly comparable to the unemployment growth that actually occurred during the Great Recession itself; they interpret this result as an explanation for the slow recovery in the unemployment rate in the years after the trough of the Great Recession. HMM study the 2014 expiration of EUC and find that that expiration was responsible for the creation of approximately two million jobs. This effect would translate into a 1.1% decrease in employment as a result of the expansion of UI from 26 to 99 weeks, which corresponds to about one third of the employment decrease of the Great Recession as measured in our data set.

[3] This calculation is made for December 2010.

However, our results are quite different from those in HKMM and HMM, despite employing apparently similar strategies. In **Online Appendix A**, we compare our results to both HKMM and HMM and we discuss in detail what accounts for the substantial differences in our respective estimates. In summary, with respect to HKMM, we have found that a few factors explain the bulk of the difference between our two sets of results. First, our dependent variable is constructed using county-level employment data from the Quarterly Census of Employment and Wages (QCEW), which is derived from administrative filings. HKMM and HMM, in contrast, use as their primary dependent variable the county-level unemployment rate from the Bureau of Labor Statistics LAUS program, which is partially model-based. Second, we handle the dynamics of the treatment effect differently. HKMM quasi-forward difference their dependent variable, and scale up their estimate to deduce the effects of a permanent change in policy. In contrast, we use a less parametric distributed lag framework to document the dynamics of the employment response in a transparent fashion over a window spanning from a year prior to treatment to two years following treatment. This provides clear evidence on endogeneity concerns, policy anticipation, and the actual impact on employment over the two years following the policy change. We also replicate HMM and find that our replication of their estimates for the 2014 expiration of the extended benefits fall close to zero when we use the most recent LAUS data, which were substantially updated in a 2015 redesign of the LAUS estimating procedure. Additionally, in an event study specification, HMM estimate a substantial negative employment effect using QCEW data. These results seem primarily driven by their choice of auxiliary parametric assumptions—namely their use of a county-specific polynomial trend model, estimated over a long time horizon. Instead of relying on a parametric counterfactual, we show that our treatment and control units across the border exhibited parallel trends prior to the expiration, display no jump at expiration and continue in parallel fashion after expiration—implying little employment effect.

More recently, two working papers have estimated the macro effect by exploiting variations in state-level UI extensions coming from measurement error in the total unemployment rate. Coglianese (2015) uses the variation between the CPS-measured unemployment rate and a constructed unemployment rate from UI records as an arguably exogenous shifter in the maximum benefit duration. Using a conceptually similar strategy, Chodorow-Reich and Karabarbounis (2016) use the variation in benefit duration coming from the gap between real-time and subsequently revised official unemployment rates. Both Chodorow-Reich and Karabarbounis (2016) and Coglianese (2015) find very small effects of UI extensions on aggregate employment. One limitation of the measurement error based approach is that the policy changes they study are less durable than the changes we examine in this paper and thus the external validity may be more limited. However, the very different types of variation leveraged across our two sets of papers makes them complementary. Our findings are also consistent with Marinescu (2015), who finds that UI benefit extensions during the Great Recession decreased job applications but not posted vacancies, implying a modest impact of the extensions on overall job finding and unemployment rates. Finally, in their case study of Missouri, Johnston and Mas (2015) find substantially larger, negative, macro employment effects than we find in this paper. Their macro estimates are similar in size to their micro estimates. Our approach differs from their macro estimates primarily in that we aggregate across many different benefit extensions and reductions and that our analysis uses variation across border counties rather than neighboring or similar states.

The remainder of the paper is structured as follows: In Section 2, we discuss important institutional details of the unemployment insurance extensions during the Great Recession that are critical for our identification strategy. In Section 3, we discuss our data. In Section 4, we discuss the identification challenges we face in our estimation and present our methodological approaches. In Section 5, we present our empirical results. In Section 6, we compare our macro estimates of UI expansion on employment with micro-level estimates based on labor supply elasticities, and back out an implied fiscal multiplier. Finally, in Section 7, we conclude.

2 Unemployment Insurance Background

The Great Recession saw a dramatic expansion of unemployment insurance benefits in all states. In part, this expansion occurred due to policies that were put in place prior to the Great Recession. However, Congress also passed legislation extending the maximum duration of unemployment insurance. In a majority of states, maximum benefit duration increased from 26 weeks to a maximum of 99 weeks depending on the state of the local labor market. In this section, we describe these extensions and how they were rolled out across states. It is precisely these differences across states—and in particular neighboring states—which we exploit in our identification of the impact of unemployment insurance benefit duration on employment.

Extended Benefits (EB)

Historically, when not in recession, most U.S. states have provided a maximum of 26 weeks of unemployment insurance to job-losers. At the onset of the Great Recession, in 2008, only two states offered more than 26 weeks of regular benefits. Massachusetts had a maximum of 30 weeks of UI benefits and Montana had a maximum of 28 weeks and no states offered less than 26 weeks.[4]

However, since Congress created the Extended Benefits (EB) program in 1970, maximum benefit lengths increase automatically when unemployment is high and growing. At a minimum, in states where the Insured

[4]Not all claimants are eligible for the maximum number of weeks of benefits. In most states, individuals with relatively weak recent labor force attachment are eligible only for a fraction of the maximum weeks of benefits. Throughout this paper, we abstract from this complication by focusing on the maximum UI duration. Our estimates, therefore, can be seen as an intention to treat effect. Johnston and Mas (2015), using micro-data from Missouri, find that approximately 70% of UI claimants had sufficient labor force attachment to be eligible for the full 26 weeks of regular benefits from 2003-2013.

Unemployment Rate (IUR) exceeds 5%, and the IUR is at least 1.2 times the IUR in the previous two years, claimants are eligible for 13 additional weeks of UI after the expiration of regular benefits.[5] The same law also provides two optional "triggers," which can be adopted by states at their own discretion. The first trigger provides for 13 weeks of EB for states whose IUR exceeds 6% (regardless of the change in the IUR over time). The other optional trigger is based on the Total Unemployment Rate (TUR): the trigger provides for 13 weeks of EB when both (1) the TUR exceeds 6.5% and (2) the current TUR is at least 1.1 times its value in the prior two years. States adopting this second trigger must provide 20 weeks of EB when (1) the TUR exceeds 8%, subject to the same growth-over-time requirement.[6] States can adopt zero, one, or both optional triggers, but no more than one trigger can be "on" at any point in time, meaning that the number of weeks of EB is capped at 20.

Normally, the costs of EB are shared equally between the federal and state governments. As a result, many states did not have statutes activating the optional EB triggers at the onset of the Great Recession. However, after the passage of the American Recovery and Reinvestment Act (ARRA), the federal government paid for the full amount of EB extensions. Some states (mostly deeply conservative ones) nonetheless declined to activate the optional triggers. For example, while Mississippi had a TUR of well over 8% continuously from January 2009 through October 2016, peaking at over 11% in 2010, they were never eligible for EB because the insured unemployment rate never went above 5.6% and the state declined to enact the optional triggers. Thus, different states had different numbers of weeks of EB in part due to differences in the state unemployment rates and in part due to state policy differences. The federal government maintained its full support of EB until the end of 2013 when it returned to the default equal cost sharing rule.

Emergency Unemployment Compensation (EUC)

In response to the first signs of a weakening labor market, on June 30, 2008, Congress and President Bush created the Emergency Unemployment Compensation (EUC) program. At first, EUC provided for 13 additional weeks of benefits for all UI-eligible unemployed workers.[7] The Unemployment Compensation Extension Act of 2008 was then signed into law by President Bush on November 21, 2008. It augmented the

[5] The Insured Unemployment Rate (IUR) is, roughly, the ratio of current regular UI claimants to the number of UI-covered jobs. The Total Unemployment Rate (TUR) is the usual "unemployment rate": i.e., the ratio of unemployed persons to persons in the labor force.

[6] From December 2010 through the end of 2013 (a period in which the unemployment rate remained high but was generally not growing), states were allowed to apply a three-year lookback period instead of a two-year lookback period for the purpose of determining growth over time.

[7] To be more precise, this legislation—and all subsequent legislation related to EUC—provided for increases in benefit lengths equal to the lesser of (1) a specified number of weeks or (2) a fraction of the number of weeks of regular benefits. For the initial legislation in June 2008, the specified number of weeks was 13 and the fraction of the number of weeks of regular benefits was 50%. For the vast majority of states that had regular benefits greater than or equal to 26, the specified number of weeks was the binding factor. For those states with fewer than 26 weeks of regular benefits, the percentage of regular benefits was always binding. In this paper, we code the weeks available under EUC exactly as specified in the law; however, in the discussion that follows, we discuss only the specified number of weeks, which applies to states with at least 26 weeks of regular benefits.

EUC program while also creating the first differences across states in their access to the EUC extensions. It authorized 20 weeks of EUC for all states (an increase from 13) and an additional 13 weeks for those with a total unemployment rate exceeding 6%.[8] These additional weeks were organized into "tiers": Tier 1 corresponded to the first 20 weeks of EUC, while Tier 2 corresponded to the baseline 20 weeks plus an additional 13 weeks. During this period, a state with 26 weeks of regular benefits could qualify for up to 79 weeks total of benefits. Then, on November 6, 2009, the Worker, Homeowner, and Business Act of 2009 further increased maximum UI duration. Tier 1 remained in place. However, Tier 2 was increased from 13 to 14 weeks and extended to all 50 states. The law also added Tier 3, providing 13 additional weeks to states with a TUR of greater than 6%, and Tier 4, providing 6 additional weeks for states with a TUR of greater than 8.5%. After the passage of this law, states had access to a maximum of 99 weeks of benefits. This schedule remained in place, with the exception of temporary lapses, until early 2012, when Congress enacted laws that slowly began to phase out EUC.[9]

On February 22, 2012, Congress passed and the President signed The Middle Class Tax Relief and Job Creation Act of 2012 which slightly lowered the generosity of the EUC in a gradual way, first starting on May 27, 2012, and then again on September 2, 2012. By September 2, 2012, Tier 1 had been scaled back to 14 weeks and was still available to all states. Tier 2 remained at 14 weeks but again became available only to states with a TUR of greater than 6%. Tier 3 was scaled back from 13 to 9 weeks and the state TUR threshold was raised to 7%. Finally, Tier 4 was increased to provide 10 extra weeks for states with a TUR of above 9%. The program finally came to an end at the end of December 2013.[10] In total, over the Great Recession, individuals in qualifying states received up to 99 weeks of unemployment insurance. Compared to the baseline of 26 weeks, this is an increase of 73 weeks; so the maximum UI benefit duration in some qualifying states increased by almost 300%.

Changes in State-Level Regular Benefits

In addition to changes in federal policy and changes in state unemployment rates which triggered changes in unemployment benefit generosity, during our sample period, UI duration was also influenced by state-

[8] A state could also have become eligible for 33 weeks with a sufficiently high IUR; in practice, the IUR trigger was never binding.

[9] There were four lapses in EUC that occurred in 2010, arising due to political disagreements regarding the extension of the program. The longest such lapse lasted from May 30, 2010 to July 18, 2010. In each of the lapses, beneficiaries were paid retroactively for any weeks of missed payments. Furthermore, during these lapses, the funding rules for EB reverted to their pre-ARRA levels, which led many states to suspend EB payments during these lapses as well.

[10] Upon the expiration of EUC at the end of 2013, EUC beneficiaries immediately stopped receiving benefit payments. Prior to the final expiration, however, the phase-out was more gradual. If a state "triggered-off" a certain tier, people who had already qualified for a given tier were allowed to finish that tier. However, beneficiaries were not allowed to move to the next tier. One exception, discussed in the following subsection, is North Carolina, which lost access to all EUC money as of July 1, 2013. In our econometric specifications, our duration variable is the maximum duration available in a given month for a new entrant into unemployment. Thus, we do not distinguish between gradual phase-outs and sudden benefit cessations.

level policy changes. Starting in 2011, some states began to lower maximum duration for regular state-level benefits below the usual 26 weeks. Arkansas reduced its maximum benefit duration to 25 weeks and both Missouri and South Carolina to 20 weeks in 2011. Then, in 2012, Florida, Georgia, Illinois and Michigan reduced their maximum benefit duration. Michigan lowered it to 20 weeks, while the other three made it contingent on the state unemployment rate. North Carolina also reduced its regular benefits to 20 weeks; additionally, North Carolina reduced the weekly benefit amount from $535 to $350, which violated its agreement with the Department of Labor. For this reason, all EUC benefits immediately expired in North Carolina, which caused its maximum benefit duration to fall by 53 weeks. The duration of regular benefits fell further in North Carolina in 2014, as it was also set to be contingent on the state unemployment rate.

Variation Between Neighboring States

Importantly, the path of benefit extensions—from regular benefits, EB, or EUC—often differed markedly across neighboring states. These differences across neighboring states were largely a result of differences in state unemployment rates, but also to some degree due to variations in state policy. It is precisely these time-varying differences across neighboring states that we use for our identification strategy. In **Figure 1**, we graphically show the evolution of the benefit generosity over time nationally, which strongly (negatively) co-varies with the national employment-to-population ratio.[11] In **Figure 2**, we show the differences across neighboring counties in the numbers of weeks of available unemployment insurance, where the reported difference is between "high" and "low" benefit duration counties, defined by comparing the average duration in the treatment period (2008m11-2013m12) versus the the prior 12 months (2007m11-2008m10) when these differences were typically zero or very small. Prior to November 2008, most counties had access to an identical amount of unemployment insurance, with the exception of those in Massachusetts and Montana. Afterwards, however, some neighboring states (and thus neighboring counties across state borders) started offering different lengths of maximum benefit duration. The average gap between states with longer versus shorter total duration within the county pairs rose to nearly 12 weeks by late 2011, before declining to an average gap of near zero with the expiration of EUC in December 2013. This variation over time is used in our full panel estimates. We also use the national level policy variation due to the the November 2008 expansion, and the late 2013 expiration, of the EUC program as instruments for our IV strategy. In **Figure 3**, we show a map of the counties that had different generosity levels right before the EUC expiration in December 2013. **Appendix Figure B1** shows the analogous map for the variation created by expansion of the EUC program in November 2008.

[11] Our measure of EPOP is below the US DOL measure. This is largely because our measure is based upon UI employment, and thus excludes those in the informal sector as well as the self-employed. Additionally, we calculate EPOP by dividing employment by the 15+ population in the county, rather than the 16+ population used by the DOL.

3 Data

We use county-level employment data from the Quarterly Census of Employment and Wages (QCEW). The QCEW data is based on ES-202 filings that nearly all establishments are required to file quarterly with their state government, for the purpose of calculating UI-related payroll taxes. These employment and earnings counts are shared by the states with the Bureau of Labor Statistics, which releases the data at the county-industry-month level. Since 98% of jobs are covered by unemployment insurance, these payroll counts constitute a near census of employment and earnings. There are some limitations: the QCEW does not capture workers in the informal sector or the self-employed, and it misses the small number of workers who participate in their own unemployment insurance system, such as railroad workers and workers at religiously-affiliated schools. Importantly, the QCEW covers both private and public sector employment.[12] The QCEW provides total employment for each month at the county level. In our baseline estimation, we require that each county be in the data set in every month. This excludes four counties for which there is at least one month in the sample where the QCEW does not report data due to confidentiality problems with disclosure. This occurs only in counties with very low population. In our robustness section, we additionally report estimates using the full unbalanced panel.

We divide employment by population of those 15 and older, which we obtain from the census at the annual level and interpolate log-linearly within each year. Prior to estimation, we seasonally adjust our dependent variables by subtracting off the county-month specific mean of the variable in question, where this mean is calculated over the period 1998-2004.[13] As we show later in the paper, however, our results are robust to using raw rather than seasonally adjusted data.

Our data on the number of weeks of regular benefits comes from Department of Labor reports which are issued biannually.[14] To account for occasional changes in the numbers of weeks of regular benefits that occur during the intervening period, we augment these data with online searches of news media and state government websites. We obtain information on EUC and EB from the trigger reports released by the Department of Labor, available at http://www.oui.doleta.gov/unemploy/claims_arch.asp. These reports provide the number of weeks of EB and tiers of EUC available for each state, in each week. When a change in weeks of benefits happens within a month, we assign the time-weighted average of the maximum duration to that month.

As discussed above, there were several lapses in the EUC program during 2010. In the popular press,

[12] We focus our analysis on total employment (the sum of private and public sector employment), though we do provide results on private employment as a robustness check.

[13] For the sake of summary statistics and the small number of specifications we estimate without county fixed effects, we add back the overall mean level of EPOP for each county measured over the 1998-2004 period.

[14] http://www.unemploymentinsurance.doleta.gov/unemploy/statelaws.asp

expectations were that these lapses would be reversed, and that the original EUC benefit durations would be reinstated. This is in fact what did happen. In our baseline specifications, we treat the lapses as true expirations—that is, those county-by-month observations are coded as having EUC equal to zero. However, we show in robustness checks that our estimates are not substantially affected if we code the benefit durations for these few months as having remained unchanged at their pre-lapse level.

We also use a list of all contiguous county pairs that straddle state borders; this data comes from Dube et al. (2010). In our baseline specifications, we have a total of 1,161 county-pairs.

In addition, we obtain county level unemployment and employment data at the quarterly level from the Local Area Unemployment Statistics (LAUS) published by the Bureau of Labor Statistics. We obtained the most current data (as of November 10, 2016) via `http://download.bls.gov/pub/time.series/la/`. We additionally obtain a vintage series of county unemployment rates and employment (prior to the March 2015 redesign) via FRED. This is the main data source used by HKMM and HMM, and we use it as part of our reconciliation exercise in **Online Appendix A**.

4 Research Design

4.1 The Identification Problem

To credibly estimate the effect of UI extensions on aggregate employment, we need to address a serious problem of reverse causality. Negative employment shocks that raised the unemployment rates were likely to mechanically raise the maximum benefit duration within the policy environment during the Great Recession. **Figure 1** illustrates the identification problem facing researchers when estimating the effect of UI extensions on employment. Between 2008 and 2014, we see a U-shaped time path of maximum benefit duration, along with an inverted-U shaped time path for the employment to population ratio.[15] However, it would be naive to assume that this correlation is causal in nature. A closer look confirms that the decline in employment in 2008 preceded the EB and the EUC tier extensions. Similarly, employment was already on the mend well before the 2014 EUC expiration occurred. It is possible that UI extensions were responsible for some of the decline and some of the persistence in the high unemployment rates the U.S. experienced in the 2009-2013 period. However, as **Figure 1** highlights, it is likely that some or much of this relationship reflects a mechanical endogeneity of UI maximum benefit duration to the state of the economy.

While the endogeneity problem is most obvious when considering time series variation, a differences in differences (or the classic two-way fixed effects) strategy is unlikely to eliminate the endogeneity bias. On

[15]To be consistent with our baseline regression specifications, this figure shows the time series of EPOP and duration taken as an unweighted average of counties.

the one hand, there was a substantial amount of variation in UI generosity over time and differentially across US states, making it feasible to use panel variation in UI duration. However, the assumption that states which saw larger increases in the maximum benefit duration had parallel employment trends with states which experienced smaller increases is unlikely to hold due to the mechanical endogeneity: the rules of EUC and EB provide for longer benefits in a given state when the unemployment rate in that state is higher. Locations which switch into offering higher benefit duration will likely be locations in decline, and locations that switch into offering lower benefit duration will be locations in recovery—likely causing a bias in the two-way fixed effects estimate.

We explicitly demonstrate the scope of this endogeneity problem by showing how high-treatment counties—i.e., counties that would eventually experience a large increase in the maximum benefit duration—had very different employment trends prior to treatment as compared to other counties. For this exercise, we construct a time-invariant, continuous measure of the average treatment intensity for each county, $treat_c$. This is defined as the difference in time-averaged maximum benefit duration in a given county during the "treatment period" (i.e., between November 2008 and December 2013) versus the 12 months prior (i.e., between November 2007 and October 2008).[16] For example, if a state's average maximum UI duration during the treatment period was 90 weeks, and the average maximum benefit length in the 12 non-treatment months was 30 weeks, it would have a value of $treat_c$ equal to 60 weeks. For ease of interpretation, we rescale this variable by dividing it by 10, so that a value of 1 corresponds to a difference of 10 weeks of treatment, which is roughly equal to the mean difference in duration between neighboring counties which straddle state borders during the treatment period. We then estimate the following model over the 2004m11-2008m10 period, i.e., the four years preceding the introduction of differential UI benefits:

$$E_{ct} = \alpha \times treat_c \times t + \lambda_c + \theta_t + \epsilon_{ct} \tag{1}$$

where t is time measured in months divided by 48.[17] λ_c is a set of county fixed effects, while θ_t is a set of common period fixed effects. Our estimate of α thus measures the difference in the linear employment trend between high- and low-treatment counties prior to November of 2008. For this specification, we cluster our standard errors at the the state level. The first column of **Table 1** shows our estimate for $\hat{\alpha}$. The estimate, significant at the 1% level, implies that EPOP declined by 0.78 percentage points in the four years prior to November 2008 in counties that would subsequently receive an additional 10 weeks of benefits. This result is consistent with the mechanical endogeneity problem discussed above, and casts doubt on the assumption

[16] This "non-treatment" value will in general not be equal to 26, since it includes the period from July to October 2008 when all states were eligible for 13 weeks of EUC.

[17] Note that there are 48 months in this sample, so the date variable equals (essentially) zero at the start of the sample and one at the end.

of parallel trends across counties prior to increases in benefit duration.[18]

4.2 Border county pair strategy

The failure of the two-way fixed effects strategy motivates us to restrict our sample to contiguous county pairs which straddle state borders (Dube et al., 2010, 2016) and estimate the effects within border county pairs. The main idea behind this strategy is that neighboring counties in adjacent states are reasonably well matched. Dube et al. (2016) show that adjacent county pairs straddling state borders are much more alike in terms of levels and trends in covariates than are randomly matched pairs of counties. However, while adjacent counties are likely to face similar economic shocks as each other, their UI maximum benefit durations will be driven by their respective states' unemployment rates and policy choices—which may be quite different. Therefore, by focusing on comparisons between border counties, we are able to account for all confounders that vary smoothly geographically, and better account for the mechanical endogeneity problem that plagues the two-way fixed effects approach. **Table 2** shows that the treated and control counties were quite similar: pre-existing characteristics seem relatively balanced between the high-treatment and low-treatment counties within pairs.

For each month t, our border county pairs (BCP) data is organized to have two observations in each pair p—one for each county c of the pair. Note that this also means that a given county c appears in the data k times (for each month t) if it borders k counties in adjacent states. Before describing in detail our key empirical specifications, we first use this BCP data to show that within-pair variation dramatically reduces the problem of pre-existing trends. We re-estimate a regression of EPOP on the time-invariant average treatment intensity, $treat_c$, and county fixed effects, similar to **Equation (1)**. But now, instead of a single set of period effects, we include a full set of pair-period fixed effects, ν_{pt}. This sweeps out the variation between pairs, and only uses within-pair variation to identify α.[19]

$$E_{cpt} = \alpha \times treat_c \times t + \lambda_c + \nu_{pt} + \epsilon_{cpt} \tag{2}$$

As before, the estimation period runs from November 2004 to October 2008. The coefficient α has a similar interpretation as in the prior strategy, but now measures the differential pre-existing employment trends by treatment status within each adjacent county pair. The results in Column 3 of **Table 1** show that for the sample of border counties, the differential pre-existing trend within county pairs (-0.24) is much closer to zero and statistically insignificant, in contrast to the estimates from the two-way fixed effects model using

[18] We show results from a two way fixed effects model in Appendix Table B2.

[19] With two observations within each pair-period group, this approach gives the identical coefficients as if we dropped the pair-period fixed effects and instead (1) took the spatial difference of the dependent variable and main independent variable across each county pair p at each time t, and (2) replaced county fixed effects by pair fixed effects.

the same sample (-0.98). This constitutes very clear evidence that the estimates using neighboring counties as controls are likely to exhibit less bias than those from the two-way fixed effects model. Moreover, the standard error from the BCP model (0.29) is not dramatically larger than that of the two-way fixed effects model (0.21), suggesting that it is a reduction in bias and not statistical power that drives the changes in statistical significance in **Table 1**.[20]

While the evidence on pre-existing trends from **Table 1** show that the BCP strategy is a very important improvement over the two-way fixed effects model, we may worry about remaining endogeneity bias, especially given the explicit reverse causality in this context. This motivates us to implement an additional data-driven refinement to the BCP strategy. In particular, we drop the quartile of pairs with the largest absolute differences in pre-existing EPOP trends over the 2004m11-2008m10 period. These BCPs appear to be more poorly matched in that the counties in these pairs exhibit qualitatively different trajectories prior to the UI extensions, and these trajectories may be mechanically correlated with subsequent UI duration.[21] Hereafter, we refer to this specification as trimming our sample based on pre-treatment trends, or PTT-trimming. Column 4 of **Table 1** shows $\hat{\alpha}$ for the PTT-trimmed sample and confirms that removing the worst-fitting quartile further reduces the extent of pre-existing trends to -0.11.

In this paper, we report estimates using several different types of regressions. First, to visually show how employment evolves on the high-treatment versus low-treatment sides of the border, we estimate a model using the same time-invariant average treatment intensity, $treat_c$, that we used above for the assessment of pre-existing trends. We regress EPOP on a a set of interactions $treat_c \times \mathbb{1}\{t = s\}$ variables, where $\mathbb{1}\{t = s\}$ is an indicator for date s. In the full sample, we omit the variable corresponding to October 2008. We additionally control for county fixed effects λ_c and pair-period effects ν_{pt}. The estimating equation is as follows:

$$E_{cpt} = \sum_{s=\tau_A}^{\tau_B} \beta_s treat_c \mathbb{1}\{t = s\} + \lambda_c + \nu_{pt} + \epsilon_{cpt} \tag{3}$$

Since $treat_c$ is a continuous, time-invariant measure, the coefficients β_s trace out how EPOP evolves in the treated versus control sides over time, as compared to a base period of October 2008, the month before the first cross-state variation in federal UI benefits in our sample.

While the time-invariant treatment measure is useful for a qualitative, visual assessment of how employ-

[20]This evidence is complementary with the evidence provided in Section 4.3 of HKMM. HKMM find substantially larger estimates of the effect of UI on unemployment when their border pair sample is replaced by a "scrambled border pair" sample, in which pairs are formed randomly (rather than by reason of geographical adjacency). HKMM argue (and we agree) that this is indicative of the role played by the BCP strategy in reducing mechanical endogeneity.

[21]Even if economic conditions evolve continuously across state borders, the statistics for a given border county will measure an average of economic conditions some positive distance away from the border. This might be a concern for geographically large counties in the western United States. In our robustness section, we show that dropping pairs whose centroids are more than 100 km apart has little effect on our estimates.

ment evolved on the two sides of the border, it does not use the timing of policy changes with any precision. Our baseline BCP-FE specification equation uses a normalized maximum benefit duration (in weeks), D_{ct}, to estimate the following equation:

$$E_{cpt} = \beta D_{ct} + \lambda_c + \nu_{pt} + \eta_{cpt} \tag{4}$$

We normalize D_{ct} by dividing the maximum benefit duration by 73, to make β interpretable as the change in EPOP from the median expansion in duration that took place in the Great Recession.[22] Again, we include county fixed effects λ_c to account for persistent differences between the two members of the pair,[23] and pair-period effects ν_{pt} to sweep out between-pair variation. Clearly, this strategy still relies on D_{ct} being uncorrelated with η_{cpt}, i.e., $E(D_{ct}\eta_{cpt}) = 0$, but now this assumption needs to hold only within a local area that is likely to be experiencing more similar economic shocks. The third column of **Table 1** shows why we believe this assumption is closer to the truth in the county-pair setting relative to the two-way fixed effects setting. **Equation (4)** is estimated for both the baseline sample of all border county pairs, as well as the PTT-trimmed sample of county pairs. The baseline regression is estimated over the period from November 2007 to December 2014, which includes the period of differential EUC (November 2008 - December 2013) as well as 12 months prior and 12 months after.

We also present the dynamics of employment around the time of the policy change. There are two specific aims that underlie this analysis. First, we wish to use the leading coefficients to detect pre-existing trends and assess the validity of the research design. Second, we wish to assess possible anticipation or lagged effects of the policy. To this end, we utilize a first-differenced distributed lags specification with a set of 11 monthly leads and 24 monthly lags, along with the contemporaneous benefit duration, D_{ct}. This specification allows us to focus on employment changes within the 36 month window around the time of treatment.

Our estimating equation for the dynamic specification is:

$$\Delta E_{ct} = \sum_{k=-11}^{24} \beta_k \Delta D_{c,t-k} + \nu_{pt} + \epsilon_{cpt} \tag{5}$$

Successively summing the coefficients traces out the cumulative response to a one-time, permanent unit change in D: $\rho_\tau = \sum_{k=-11}^{\tau} \beta_k$ represents the cumulative response at event time, τ.[24] For ease of interpretation, we center the cumulative responses around a baseline of the month just prior to treatment,

[22] All but two states had 26 weeks of benefits prior to the onset of the Great Recession, and the median as well as mode for state UI duration was 99 weeks from November 2009 until April 2011.

[23] We replace county fixed effects with county-cross-county-pair fixed effects in the small number of specifications in which the panel is unbalanced.

[24] Note that β_k is the response associated with D_{t-k}. This indexation convention allows us to index the coefficients by event time.

$\tilde{\rho}_\tau = \rho_\tau - \rho_{-1}$, which imposes that $\tilde{\rho}_{-1} = 0$. We plot the centered cumulative response $\tilde{\rho}_\tau$ by event time, along with the associated confidence intervals below.

While the border county pairs strategy provides greater internal validity, one potential concern is about the representativeness of border counties. Summary statistics in **Appendix Table B1** confirm that border counties are relatively comparable to the full set of counties, indicating that the sample restriction for purposes of internal validity comes at minimal sacrifice of external validity.

4.3 Instrumental variables estimation: EUC Policy Changes

Estimating **Equation (4)** by OLS exploits all of the variation in maximum benefit duration induced by both policy changes (EUC, state adoption of optional EB triggers, and state changes to regular benefits) and endogenous movements in state unemployment rates across various thresholds (from EUC and EB triggers). That is, our OLS specification has the undesirable feature that it exploits variation in benefit duration in a given month which was caused by a change in contemporaneous state-level unemployment. By only comparing adjacent border counties, we are likely to reduce the scope of the endogeneity problem, since the employment shocks affecting policy are from the state as a whole, while we are accounting for the county's employment shock by comparing it to its cross-state neighbor. Nonetheless, to the extent that endogeneity bias may remain, we can further reduce it by restricting the variation we use to national-level policy changes. Counties within a border pair are less likely to have systematically different employment trends when UI duration changes due to national policy than when one county's state is triggering on or off of EB or an EUC tier. We therefore develop an instrumental variables approach that isolates the effects of cross-border changes in benefit duration that are triggered by persistent changes in national policy, and not by contemporaneous economic shocks.

The first policy change that we use is the passage of the Unemployment Compensation Extension Act (UCEA) in November of 2008, which granted states 20 weeks of federally funded benefits, or 33 if the total unemployment rate at the time exceeded 6%. This led to an increase in UI benefit durations which varied across states, introducing the first across-state variation in EUC availability in our sample.[25] The second national policy change we use is the expiration of the EUC program in December 2013, which led to a larger reduction in UI duration which also varied across states.

Of course, the change in national policy creates variation precisely because there were differences in the *level* of unemployment across states. For the 2008 policy change, states that had a TUR exceeding 6% saw a bigger increase in benefit duration than states with a lower TUR. Similarly, for the 2014 expiration, states

[25] Prior to UCEA, variation in federally provided benefits existed in two states: North Carolina and Rhode Island were eligible for 13 and 20 weeks of EB, respectively, at the time of the policy change. No other state was eligible for EB at that time.

with higher unemployment rates experienced larger reductions in benefits. While high and low unemployment states very well may have been on different trajectories around these two events, the BCP strategy is arguably better able to account for such trends compared to times when the policy change is directly induced by changes in state unemployment rates.

For our IV specifications, we use a two year window—one year on each side of the national policy change. We regress EPOP on weeks of benefits, controlling for pair-period fixed effects and county fixed effects. We then instrument benefit duration with a variable that reflects only the change in duration caused by the EUC policy change. The instrument does not exploit variation caused by EB triggerings, EUC triggerings, and state-level policy changes. Our two stage least squares estimation strategy is thus given by the set of equations:

$$E_{cpt} = \beta D_{ct} + \lambda_c + \nu_{pt} + \eta_{cpt} \tag{6}$$

$$D_{ct} = \beta_z z_{ct} + \rho_c + \gamma_{pt} + \epsilon_{cpt} \tag{7}$$

where the instrument z_{ct} reflects the instantaneous change in the maximum UI duration available in the county due to the national EUC policy change. The instrument z_{ct} is defined as follows:

$$z_{ct} = \begin{cases} D_c^{08} & \text{Nov. 2007 - Oct. 2008} \\ D_c^{08} + \delta_c^{08} & \text{Nov. 2008 - Oct. 2009} \\ D_c^{13} & \text{Jan. 2013 - Dec. 2013} \\ D_c^{13} - \delta_c^{13} & \text{Jan. 2014 - Dec. 2014} \end{cases}$$

For the 12 months prior to the 2014 policy change, we set the value of z_{ct} to equal the number of weeks of UI available in the last week of December 2013 (immediately prior to the EUC expiration), D_c^{13}. For the remaining 12 months in the sample, we subtract from D_c^{13} the number of weeks of benefits lost as a result of the EUC expiration $\left(\delta_c^{13}\right)$, and set z_{ct} equal to this value.[26] For the two year window around the 2008 policy change, the instrument is defined analogously, using the maximum UI duration available just before $\left(D_c^{08}\right)$ and just after the introduction of the new EUC program. Therefore, the jump in z_{ct} that occurs in November 2008 $\left(\delta_c^{08}\right)$ exactly equals the differential number of weeks made available by the onset of the UCEA. We also pool both events together, and estimate this model using the 24 months of data around the

[26]Therefore, the change in the instrument z_{ct} between December 2013 and January 2014 takes into account the decline in duration explicitly resulting from the EUC expiration, but not any contemporaneous changes in state-level regular benefits. In our robustness section, we show results from a specification where the instrument also takes into account the five state-level policy changes that occurred at the same time as the national policy change.

2008 onset along with the 24 months of data around the 2014 expiration.[27] For all of these specifications, we estimate the results using the complete baseline BCP sample as well as the refined (PTT-trimmed) sample. Because the EUC program in North Carolina expired at the end of June 2013 (rather than December), we drop county pairs that include a North Carolina county from the 2014 subsample in the baseline analysis.[28]

We additionally show reduced form and first stage estimates underlying the IV regressions by month relative to the event. As with OLS, the dynamic specification is estimated in first differences:[29]

$$\Delta E_{cpt} = \sum_{\tau=-12}^{11} \beta_\tau \delta_{ct} \mathbb{1}\{eventdate_t = \tau\} + \nu_{pt} + \eta_{cpt} \tag{8}$$

$$\Delta D_{ct} = \sum_{\tau=-12}^{11} \beta_{z\tau} \delta_{ct} \mathbb{1}\{eventdate_t = \tau\} + \gamma_{pt} + \epsilon_{cpt} \tag{9}$$

We define $\delta_{ct} = \delta_c^{08}$ for the 2007-2008 sample and $-\delta_c^{13}$ for the 2013-2014 sample, each divided by 10 for the ease of interpretation. As with OLS, the sum of coefficients $\rho_\tau = \sum_{k=-11}^{\tau} \beta_k$ and $\rho_{z\tau} = \sum_{k=-11}^{\tau} \beta_{zk}$ represent the cumulative response by event time. These represent the average within-pair differences in employment and the prevailing maximum benefit duration—over a 24 month window around the national policy change—for a pair in which the difference in the instantaneous increase in maximum benefit duration (due to the policy change) was 10 weeks. We omit the variable corresponding to $eventdate_t = -1$ (which corresponds to October 2008 and December 2013), meaning that the plotted coefficients are centered relative to date -1 leading values.[30]

It is useful to consider where our policy variation is coming from when using this IV approach along with BCP sample. Consider two adjacent counties, A and B, which followed similar employment trends prior to October 2008, but where side A saw a larger increase in benefit duration in October 2008 because it happened to be in a state with an already high state unemployment rate. Variation in policy, then, is coming largely from more negative *past* employment shocks in *other* counties in A's state—as compared to past employment shocks in other counties in county B's state. The same logic applies to EUC expiration in December 2014. The combination of more plausibly exogenous variation due to national policy changes with local cross-state comparisons guards against endogeneity bias by putting both geographic and temporal

[27]For this pooled specification, we allow the county fixed effects to vary across the two subsamples (that is, the county fixed effects are replaced with county-by-subsample fixed effects).

[28]In the robustness section, we report results from specifications which keep North Carolina as well as others which redefine the instrument for North Carolina to exploit variation from its earlier benefit cut.

[29]We note that estimating this model in levels (i.e., using E_{cpt} and D_{ct} and mean differencing) versus first-differences is immaterial in this case where we are estimating monthly coefficients, β_τ, over a fixed 24 month sample. Estimating the model in levels yields numerically identical estimates.

[30]For ease of interpretation, we omit January 2014 instead of December 2013 in the first stage when constructing the graph that analyzes only the 2014 expiration event. This allows the graph to show a drop in relative benefits roughly from 10 to 0 rather than 0 to -10. As we do not report standard errors for this specification, this amounts to a simple vertical shift of the graph.

distance between shocks in employment in a county and the shocks that drive the policy.

4.4 Standard errors

Except where noted, our standard errors are clustered two-way at the state-pair level and at the state level. Clustering at the state-pair level is designed to account for common, serially correlated shocks to local economies. We also cluster at the state level to account for the mechanical correlation in error terms that is introduced when one county borders counties in at least two states (and thus appears in multiple state-pairs) as well as any state level shocks. Note that our clustering strategy fully accounts for the appearance of a single state multiple times in the border county pair sample.

5 Empirical Findings

5.1 Motivating graphical evidence

Figure 4 plots the regression coefficients for the time-invariant average treatment intensity measure, $treat_c$, period by period, using **Equation (3)**. The figure plots two sets of coefficients: one with EPOP as the dependent variable, and the other with maximum UI benefit duration as the outcome. This figure shows that the side of the border receiving a larger treatment (averaged over the full treatment period) experienced a slight decline in employment starting several years prior to treatment, though this pre-existing trend is not statistically significant. The differential employment trend greatly accelerated between 2009 and 2012—at a time when the UI extensions are implemented. This might indicate a causal effect of the UI extensions. However, contrary to that interpretation, employment continued to fall at a similar rate on the side receiving a larger treatment in the post-2011 period when UI generosity difference within the pair was in decline.

Figure 5 shows the results of the same analysis using our refined PTT-trimmed set of border county pairs, where we exclude the pairs with the largest differences in pre-existing trends. The findings are reinforced when we consider this refined BCP strategy. Over the 2004-2014 period, employment on the side of the border receiving greater treatment remained essentially unchanged, even as benefit duration rose sharply in late 2008, and then dropped sharply in late 2013. This figure provides compelling visual evidence of the validity of the refined BCP design (no pre-existing trends), and that any causal employment effect of of the policy is likely to be quite small.

Together, the two figures convey several important features of the data and the research design. First, when using the baseline BCP sample, the monotonic decline in employment on the high-treatment side of the border throughout the entire period—both when UI benefit duration difference within the pair is increasing and when it is decreasing—previews our regression results that overall employment effects are likely to be modest in that specification as well. Second, trimming on pre-treatment trends eliminates not only trends prior to treatment but also the secular decline in EPOP post-treatment. These findings suggest that the secular employment decline was due to poor match quality in a minority of observations rather than a causal effect of treatment.

5.2 Main Estimates

We present our full-sample OLS estimates for the time period from November 2007 to December 2014 in the top panel of **Table 3**. This panel reports two columns of regressions estimating **Equation (4)**. The first column reports results using the baseline (i.e., untrimmed) BCP sample and the second column reports results using the sample that we refined based on pre-existing trends (the PTT-trimmed sample). The point estimate for the baseline BCP sample is 0.430. Recall that we normalized D by dividing the maximum benefit duration by 73 weeks, so this allows us to interpret the coefficient as the estimated impact on EPOP from an increase in maximum benefit duration from 26 to 99 weeks. Consequently, the baseline BCP estimate suggests that the 73 week increase in maximum benefit duration raised the EPOP ratio by 0.430 percentage points. The standard error is 0.466 and thus the estimate is not statistically distinguishable from zero. When we restrict the analysis to the PTT-trimmed sample in Column 2, the coefficient falls to 0.213. Even though the PTT-trimmed sample size is 25% smaller than the baseline BCP sample, the standard error for the PTT-trimmed estimate is smaller at 0.270: trimming on PTT rids the sample of poorly matched county pairs and thereby reduces residual variance. As a consequence, when moving from the baseline BCP to the refined BCP estimates, the maximal *negative* impact of expanding UI from 26 to 99 weeks which can be rejected at the 95% level of confidence falls in magnitude from -0.483 to -0.316.[31]

Figure 6 visually displays the employment dynamics around the treatment event in a transparent manner using the first-differenced distributed lag specification of **Equation (5)**. These estimates are useful for assessing policy anticipation and lagged effects of the policy, as well as possible biases in the research design arising from pre-existing trends. The figure shows the cumulative response in employment ($\tilde{\rho}_\tau$) starting 12 months before treatment, and extending up to 24 months after. Recall that these cumulative responses are centered at event time $\tau = -1$, so the estimates of confidence intervals for $\tilde{\rho}_\tau$ are expressed relative to the month before treatment. The top panel displays the coefficients for the full sample of BCPs, while the bottom panel displays them for the PTT-trimmed sample. For both specifications, during the twelve months prior

[31] **Appendix Table B2** presents results from the two-way fixed effects model for the all-counties sample and the border county pair sample. The point estimates are somewhat more negative, consistent with the problem of pre-existing trends documented in **Table 1**. Nonetheless, the unweighted estimates which are most comparable to **Table 3** are modest in magnitude: -0.385 (with a standard error of 0.355) for the all-counties sample, and -0.382 (with a standard error of 0.361) for the border counties sample.

to treatment, i.e., between $\tau = -12$ and -1, there is little change in employment. The leading values of the cumulative responses range between -0.321 and 0.403, and are never statistically distinguishable from zero. Overall, the distributed lag specifications produce little evidence to indicate reduced hiring in anticipation of the policy change.

Following treatment, both the baseline BCP specification and the PTT-trimmed specification show no change in employment over the 24 months following the policy change. The cumulative responses are typically positive and not statistically significantly different from zero. Even as the precision declines for longer lags, 12 months after the policy change, we can nonetheless still rule out employment effects more negative than -0.6 with 95 percent confidence for both specifications. Overall, the dynamic evidence from the OLS model suggests little employment change in the year prior to treatment (e.g., through anticipation), or during the two years following the policy change.

The instrumental variables estimates from **Equations (6) and (7)** are presented in the bottom three panels of **Table 3**. In panel 2 of **Table 3**, we report our pooled results using both the 2008 introduction (i.e., a positive treatment) and the 2014 expiration of the EUC (i.e., a negative treatment). For our preferred PTT-trimmed specification, the first stage F-statistic for the excluded instrument is 262.3, indicating that the instantaneous changes due to the national policy changes were responsible for a sizable fraction of the variation in benefit duration over the event window; the first stage coefficient is 0.842.[32] Our preferred PTT-trimmed second stage estimate is close to zero (-0.069), with a standard error of 0.635. While less precise than the OLS estimate, these estimates using only national level policy changes in the PTT-trimmed sample can rule out employment reductions of -1.31 percentage points from the 73 week expansion of maximum benefit duration during the Great Recession. The point estimate from the untrimmed BCP sample is similar (0.143), though less precise with a standard error of 0.964.

To assess the employment dynamics around the national policy changes, **Figure 7** shows the first stage and reduced form estimates period by period around the event date, as compared to the values from the month just prior to treatment (i.e., -1). The EPOP difference between the two sides of the border is plotted on the left hand Y-axis, with the difference in maximum benefit duration plotted using the right hand Y-axis. The top graph uses the baseline BCP-FE sample while the bottom graph uses the refined PTT-trimmed sample. The dynamic evidence mirrors the numerical results in **Table 3**. At date 0, there is (by construction) a clear increase of approximately 10 weeks in the maximum benefit duration relative to the neighboring county.[33] Much of this increase in benefits persists over the following 12 months. There is little

[32] If the only changes in duration in the year before and the year after policy change were due to the policy change itself, the first stage coefficient would be 1.

[33] The increase is not exactly 10 weeks because the policy changes in question did not occur precisely at the end of a calendar month.

indication of a differential trend in employment prior to the national level policy changes, which provides additional validation for the IV coupled with the border county design. Importantly, employment remains fairly stable over the 12 months following treatment and we see little indication of job loss following the national level policy changes. Furthermore, the results are visually similar both in the baseline BCP-FE and the refined PTT-trimmed sample.[34]

The pooled estimates combine both the positive treatment in 2008 and the negative treatment in 2014. We also show the disaggregated effects from each of these treatments. The 2008 results using the 2007m11 to 2009m10 period are reported in the third panel of **Table 3**, and we show the corresponding graphical evidence in **Figure 8**. Again, there is a strong first stage (the F-statistic for the excluded instrument is over 40), though this first stage is substantially weaker than the pooled first stage or the 2014 first stage discussed below. As **Figure 8** shows, the duration differences created by the implementation of UCEA in 2008 were somewhat less persistent. The more limited persistence is also reflected in the first stage coefficient of 0.729, as shown in **Table 3**. This is unsurprising given the economic turbulence and resulting triggering that followed the UCEA of November 2008.

In general, the second stage estimates from the 2008 event study are fairly noisy. The estimate on the baseline BCP sample is 0.549, with a very large standard error of 2.515. The large standard error is likely because (1) there was a lot of variability in the drop in EPOP across counties during the early part of the Great Recession, substantially increasing error variance (reduced form), and (2) the duration differences created by UCEA were less persistent (first stage). Turning to our preferred PTT-trimmed sample, the coefficient falls to 0.198, while the standard error also halves to 1.265. While the standard error remains large, the PTT-trimmed sample is somewhat more precise due to a smaller residual variance. As shown in **Figure 8**, however, there is little indication of systematic employment changes—either in the year prior to the 2008 UCEA implementation, or during the subsequent year. Overall, while noisy, the estimates from the 2008 event (especially from our preferred, more precise trimmed sample) are broadly consistent with those from the pooled estimates and do not indicate substantial losses in employment from this policy change.

Panel 4 of **Table 3** reports our IV results from the 2014 elimination of EUC. The EUC program expired at the end of December 2013, leading to large reductions in UI generosity in almost every state. Importantly, some states experienced substantially larger reductions in benefits than others. For example, benefits were reduced by 47 weeks in Illinois, Nevada, and Rhode Island, but only by 14 weeks in Virginia, Iowa, New Hampshire, Minnesota and 10 other states. **Figure 3** shows a map of the reduction of UI duration at the

[34] We also estimate the model using a sample trimmed based on trends estimated over the 2004m11-2007m10 period in order to address any concerns that PTT-trimming is mechanically eliminating anticipation effects. The graphical results, presented in **Appendix Figure 8**, are quite similar to the results presented in **Figure B2** Regression results are presented in the robustness section, below.

end of 2013. As discussed above, North Carolina lost all EUC benefits and the maximum benefit fell to 20 weeks a full six months before the national EUC expiration. As a result, we remove North Carolina from our 2014 event study sample.[35] However, in the robustness section below, we show the results from specifications in which North Carolina is included in the sample.

We show our results graphically in **Figure 9**. The figure does not show much of an effect on EPOP from the program expiration. Of note, the duration differences between county pairs were much more persistent (looking backward in time), mostly exceeding 80% of their immediate pre-expiration duration during the entirety of 2013. This explains why the first stage coefficient is much closer to unity: 0.915 for the baseline BCP sample and 0.903 for the PTT-trimmed sample. The first stage F-statistics are very high: 393 for the baseline sample and 424 for the PTT-trimmed sample. The point estimates are slightly lower than the pooled sample, at -0.024 and -0.182, respectively. However, the standard errors are substantially smaller than the 2008 analysis: 0.562 for the baseline BCP sample, and 0.521 for the PTT-trimmed sample. These estimates suggest a relatively precise null estimate of the effect of UI extensions on employment.

Although not statistically distinguishable from zero, the point estimates for the 2008 analysis are somewhat more positive than the 2014 estimates or the full sample OLS results. If these differences are real, and not merely noise, one speculative possibility is that the estimates from 2014 are less positive because they are estimated at a point in time when aggregate demand multipliers are lower.

Overall, both the OLS and IV estimates suggest that there was no sizable positive or negative employment effect of the 73 week increase in UI maximum duration during the Great Recession. This is true when we use all policy variation in our OLS specifications, or when we instrument the policy variation using national level changes. Our dynamic evidence suggests no employment changes for the first year and a half following the policy innovations. And when we consider our preferred refined BCP strategy that excludes some of the more poorly matched pairs, we find no evidence of employment changes up to 24 months following treatment.

5.3 Robustness of estimates

In this subsection, we perform a number of robustness checks. First, we show how our estimates vary with the sample period used in our estimation, and why we believe this validates our use of the refined BCP sample that trims on match quality. In the second subsection, we show how our refined PTT-trimmed results vary as we alter the trimming threshold. In the third subsection, we consider our results' robustness to a wide range of other specification choices and controls.

[35] To be clear, in the pooled estimates reported above, we include North Carolina in the 2007-2009 portion of the sample but exclude it from the 2013-2014 portion.

Choice of sample period

Table 4 shows results from the full sample OLS specification for alternative samples beginning in 2007m11, 2006m11, 2005m11, and 2004m11. The first column shows results for the baseline BCP sample and the second column shows results for the PTT-trimmed sample. Overall, the baseline BCP estimates range between 0.430 and -0.330, while the PTT-trimmed estimates range between 0.213 and 0.064. Importantly, while the estimates differ in size, we stress that none of the eight estimates shown in **Table 4** is statistically significant at conventional levels, and six of the eight are positive in sign.

At the same time, the baseline BCP estimates vary somewhat by sample, and these estimates decrease monotonically in the length of the window: the earlier the sample start date, the more negative the estimate. The gap between the estimate for the sample starting in November 2007 to the sample starting in November 2004 is non-trivial; it represents a differential impact of roughly 0.75 percentage points of EPOP from a 73-week increase in UI duration. Note that the pattern in the estimated effect is consistent with the presence of a downward trend in EPOP in treatment counties relative to control. As we discussed above, and as shown in **Figure 4**, between 2004 and 2008 we see a relative decline in EPOP on the side of the border that would eventually have higher UI duration. By pushing the start date further back in time, we are only adding data from the pre-treatment period; there is essentially no variation in UI benefits between 2004 and 2007. Adding observations from a time period when EPOP was relatively higher on the high-treatment side and when treatment was low makes the estimated treatment effect more negative. The fact that the estimated effect varies across the different sample periods leads us to believe that the baseline specification with pair-period fixed effects may reflect a degree of residual endogeneity. Put another way, a 2007m11-2014m12 sample frame − with twelve months before treatment begins and after treatment ends − ensures that any differential trends between counties is approximately orthogonal to D, our independent variable of interest. This orthogonality implies that differential trends have relatively little effect on our estimates. By contrast, with a larger amount of time before treatment than after treatment, these trends are no longer orthogonal to D, potentially leading to bias.

The variation in estimates is much smaller for the PTT-trimmed estimates: the 2007-2014 estimate is 0.213 and the 2004-2014 estimate is 0.064. We believe that this relative robustness to sample date validates the use of this refined sample (selected based on an absence of pre-treatment trends): even as the sample window becomes more asymmetric around the "treatment" period, the estimates do not change substantially, suggesting that differential trends are much smaller in magnitude in this sample. Additionally, the standard errors for the PTT-trimmed samples are also uniformly lower by between 16% to 42%, consistent with better match quality in the refined BCP sample.

Trimming on pre-treatment trends

The refined BCP strategy trims the pairs with the worst matches—25% of the sample with the biggest absolute differences in pre-treatment employment trends. In **Table 5**, we show how our four main estimates (OLS, 2008 IV, 2014 IV, and Pooled IV) vary as our threshold for trimming on PTT varies. We show estimates for different trimming thresholds across 7 rows. The rows are, respectively: no trimming, 10% trimming, 20% trimming, 25% trimming, 30% trimming, 40% trimming, and trimming at the median of the difference in PTT. The 25% trim is our main PTT specification from **Table 3**. In all four columns (i.e., for all 4 specifications), the range in the point estimates across trimming thresholds is below 1 standard error in magnitude. The coefficient estimates are fairly robust to changes in the trimming threshold. The standard error is minimized for the full sample at a 25% trim. It is minimized at a 10% trim for the pooled IV sample and the 2014 sample. It is minimized at a 30% trim for the 2008 sample. Thus, our choice of a 25% benchmark trim across all specification is a reasonable one.

Additionally, for all specifications, the primary impact of trimming seems to be a reduction in the standard errors by improving the match between high-treatment and low-treatment counties. It does not seem to systematically change the magnitude of the estimate in a positive or in a negative direction. The reduction in the standard errors is often up to 50% from the baseline sample. The one exception is the 2014 IV estimate where the maximum reduction across trimming thresholds is approximately 20%.

Other robustness checks

In **Table 6**, we consider a number of other robustness checks for our OLS estimates on the full 2007-2014 sample and for our pooled IV.[36] We do this both for the baseline BCP sample as well as the PTT-trimmed sample. The first row in the table reproduces the estimates from **Table 3**. Each of the remaining rows varies the specification, data, or sample as follows. We show estimates of impacts on private employment only. As an additional strategy to mitigate residual mechanical endogeneity, we drop pairs containing counties that show a high correlation between county EPOP and the EPOP of its state over the 2004m11-2008m10 period ("correlation trimming"). Comparison within these county pairs should be less prone to contamination from state-specific employment shocks that endogenously determine state-level benefit duration. We include an (in sample) county specific linear trend (ISLT) control. We trim based on pre-treatment trends estimated over the 2004m11-2007m10 period (instead of 2004m11-2008m10) to address concerns that PTT trimming could be mechanically removing anticipation effects. Because the lapses (correctly) might not have been seen as changes because they were expected to be reversed in a very short period of time, we recode treatment

[36]The corresponding results for the separate 2008 and 2014 IV regressions are shown in the Appendix.

during temporary lapses at the level of the duration during the last week before the lapses; we do not recode for the IV estimates because none of the lapses occur during the relevant sample periods. We also estimate using quarterly as opposed to monthly data: once using the same QCEW employment data but aggregated to the quarterly level, and once using quarterly employment statistics from a different data set, the Quarterly Workforce Indicators (QWI). We show results using data that have not been seasonally adjusted. To demonstrate that our controls are well matched to our treatments, we show robustness to restricting the sample to a plausibly better-matched group of pairs whose population centroids are less than 100 km apart. We also estimate a specification where we allow for imbalance in our panel by including counties with missing values in the sample. In addition, we show a pooled IV specification where we instrument using the total change in benefits rather than the change in benefits due solely to the expiration of EUC. In this case, the instrument includes the additional decreases below 26 weeks made by state governments in Florida, Georgia, Kansas, and South Carolina, as well as an increase from 26 to 30 weeks in Massachusetts. We also show three different specifications where we alter our baseline treatment of North Carolina, which lost access to EUC benefits earlier than other states.[37] Finally, as a further alternative, we use a log-log specification instead of the level-on-level specification used throughout the paper. We do this using both log employment and log EPOP as outcomes, but also report the EPOP-equivalent estimates in square brackets for comparability.[38]

For the OLS specifications in Columns 1 and 2, the range of the estimates from these changes is not substantial. The coefficients (or EPOP-equivalent coefficients as is the case when using logged outcomes) range between -0.145 and 0.692. In no case do we see any indication of substantial disemployment effects of the UI extensions. For the IV specifications in Columns 3 and 4, the estimates range between -0.147 and 0.930 for the baseline BCP sample, and between -0.406 and 0.659 for the PTT-trimmed sample. The greater variability for the IV is consistent with the IV estimates being more imprecise, and the standard errors are two to three times as large as the OLS counterparts. However in none of these cases are the estimates statistically distinguishable from zero.

In **Appendix Table B3**, we show the robustness checks for the 2008 and 2014 IV analyses separately. The results are largely similar to our pooled IV results, though the standard errors are significantly larger

[37]Recall that North Carolina lost access to EUC at the end of June 2013. This was a full 6 months before the other states lost access to EUC benefits, which means that North Carolina gets treated half way through the control period in the 2014 IV analysis. In our main specifications analyzing the 2014 EUC expiration, therefore, we drop all county pairs containing a county from North Carolina. We also drop North Carolina from the 2014 part of the sample in the pooled IV regression. As robustness checks, we drop North Carolina from the entire baseline BCP-FE full sample estimation as well as from the entire pooled IV specification. We also include North Carolina in the 2014 portion of the pooled IV specification. Finally, we retain the inclusion of North Carolina in the 2014 portion of the pooled IV sample but redefine the instrument, in North Carolina's case, to reflect the drop in EUC benefits for North Carolina in July 2013.

[38]For instance, the estimate of 0.006 in column 1 for log EPOP would imply that the expansion of UI from 26 to 99 weeks increased EPOP by $((\frac{99}{26})^{.006} - 1) \times 42 = 0.35$ percentage points (since the unweighted mean EPOP in this sample is approximately 42), similar to the coefficients that we see in the level-on-level specification (0.430). The level equivalents for the log-log specification are displayed in brackets below the coefficient estimates. The level-on-level equivalents of the log employment estimates are quite close to the original estimates.

for the 2008 IV and often 30-50% smaller for the 2014 IV. The 2008 IV estimates are imprecise because the initial 2008 triggering explains less of the variation in treatment in the surrounding 2 year sample period. In addition, they are imprecise because of the large variation in EPOP during the onset of the Great Recession.

5.4 External validity: size and persistence of policy changes

One potential concern with our border county pair design—or any county panel design for that matter—is whether the differences in UI benefit duration between counties across the state border were sizable and persistent, especially as compared to the national level changes in benefit duration that took place during the Great Recession. **Figure 10** shows the distribution of differences in maximum benefit duration across county pairs and over time for the full sample. Here each observation is a county pair in a given week between November 23, 2008, and December 22, 2013. As the figure shows, around 40% of pair-week observations in this sample have no difference in UI benefit durations. However, nearly half of the observations have a benefit duration exceeding 10 weeks. To put this in perspective, a 10 week differential is almost 40% of the typical maximum benefit duration of 26 weeks that prevailed in all but two states prior to the Great Recession. Therefore, the gaps across state borders that we are evaluating are economically substantial. In **Appendix Figure B3,** we show that similar sized duration gaps existed between the two sides of the border just prior to the EUC expiration in 2014.

The gaps in UI benefit duration between neighboring counties across the border were substantial, but were they also persistent? **Figure 11** shows the mean benefit duration gap (as a share of the initial gap) by weeks following a particular event.[39] On average, ten weeks after the event, 70% of the original gap in maximum benefit duration between the two sides of the border remained in place. Even 52 weeks after the event, on average, more than 50% of the original gap in duration persisted across the border. Overall, the evidence suggests that the benefit durations we are using for identification are not transitory policy shocks. The duration series in **Figures 8** and **9** show similar information for the specific 2008 and 2014 events.

We additionally show that the high average persistence of the policy shocks is not driven by a small number of cases but rather policy persistence was widespread across counties. In panel A of **Appendix Figure B4**, we show the share of counties where the duration gap continuously remained at least as large as the initial gap by weeks following the the 2008 event. The figure shows that after approximately 20 weeks, the initial gap remained in place or increased in about 60% of the county pairs; by 40 weeks, about 15%

[39] In this analysis, all changes in relative benefit differences are treated as "events" or "shocks." With the data organized at the pair-by-shock (ps) level, we regress the change in relative duration on a set of $shock_{ps} \times eventdate_{\tau}$ indicator variables, where $shock_{ps}$ is the size of the initial shock and $eventdate_{\tau}$ runs from zero to 51 weeks after the initial shock. For instance, suppose at time t, county A increased duration from 53 to 63 weeks while county B held constant at 47 weeks, then $shock_{ps}$ would be equal to 10. The dependent variable in the regression (for $\tau = 0, 1, ..., 51$) would be equal to $D_{A,t+\tau} - D_{B,,t+\tau} - 6$, since the pre-shock difference was 6 weeks. Therefore, the regression coefficients trace out the share of the original shock that remains after τ weeks.

of the pairs retained the full gap. Panel B shows evidence for the 2014 expiration. Even 50 weeks before the EUC expiration, over 40% of counties had gaps in duration at least as large as the gap at the time of expiration. Thus, the 2014 event study estimates are based on the expiration of highly persistent differentials across county pairs.

Overall, while the cross sectional differences in size and persistence of the UI benefit durations are not as dramatic as the overall national level changes that occurred during the Great Recession, they are nonetheless quite substantial—especially for the 2014 expiration event. Moreover, the persistence of the events in our samples are quite a bit greater than those used in some of the other papers in the literature. For example, the measurement error based identification used in Chodorow-Reich and Karabarbounis (2016) uses treatment events whose half life is roughly 8 weeks (see their Figure 2). In contrast, as shown in our **Figure 11**, the half life of the typical event used for our baseline OLS estimate exceeds 52 weeks.

6 Rationalizing Macro and Micro Effects of UI Extensions

A higher benefit duration has an unambiguously negative labor supply effect through increasing reservation wages. In the UI literature, the micro-based estimates of extensions on employment reflect only these labor supply considerations. In this section we compute and interpret the gap between our macro estimate and some of the prevalent micro estimates from the literature.

How do the macro effects of UI extensions on employment that we estimate compare with employment change implied by micro-level labor supply elasticities? In order to answer this question, we first express both our macro estimates and the micro literature estimates in numbers of jobs. This entails multiplying our estimates (which are in terms of EPOP) by the 15+ population in 2012 (253 million) and the micro-estimates (which are in terms of unemployment rates) by the 2012 labor force (134 million). The gap between the macro and the micro estimates of the UI extensions on employment can be written as:

$$GAP = \Delta E_{MACRO} - \Delta E_{MICRO} = (\beta_{MACRO} \times P + \beta_{MICRO} \times L)$$

where β_{MICRO} is a micro estimate from the empirical literature of the impact of raising the UI benefit duration from 26 to 99 weeks on the unemployment rate, L is the size of the labor force (in 2012), β_{MACRO} is an estimate from this paper, P is the 15+ population in 2012, ΔE_{MACRO} is the predicted change in national employment from increasing UI benefit duration from 26 to 99 weeks using our estimates, and ΔE_{MICRO} is the predicted change in national employment from increasing UI benefit duration using micro estimates from the literature. In **Table 7** we report computations using 6 estimated micro responses to the impact of increasing UI duration from 26 to 99 weeks in the literature. Five of these are from four papers estimated using data from the Great Recession (Daly et al. (2012); Farber and Valletta (2015); Johnston and Mas (2015); Rothstein (2011)). Four of these numbers range between 0.1 to 0.8. Johnston and Mas (2015) is much larger in magnitude at 4.6. We also use one estimate from before the Great Recession which comes from Elsby et al. (2010): 2.4.[40] In addition, we use two estimates of β_{MACRO} from Column 2 of **Table 3** (-0.069 and 0.213, rounded to -0.1 and 0.2 for simplicity). For each combination of estimates, we calculate the employment gap between the macro and micro employment estimates.

Our macro estimates imply a range of employment change between -0.3 million and 0.5 million. In contrast, the range implied by the micro elasticities is -6.2 million and -0.1 million; excluding the Johnston and Mas estimate, the range is -3.2 million to -0.1 million. For 11 out of the 12 combinations of estimates, the predicted macro employment change is more positive than the predicted micro change, sometimes sizably so.[41]

One explanation for a more positive macro than micro effect is the Keynesian aggregate demand channel. UI puts cash in the hands of unemployed individuals whose earnings in the absence of UI payments are likely to be well below their permanent incomes. These individuals are likely to be liquidity constrained and thus a dollar of UI expenditures is highly likely to be consumed. Empirical work has shown that the marginal propensity to consume out of one-time tax rebates during the Great Recession was 25% (Sahm et al. (2012)). Though lower income individuals responded more, the differences were not large. However, economic theory suggests that liquidity constrained unemployed individuals should have a substantially larger response to cash receipts than other groups. If UI recipients spend most of their money on consumption, this can impact aggregate demand. The total impact will depend upon the fiscal multiplier, over which there is substantial disagreement among macroeconomists. For example, when analyzing the likely impact of the ARRA, the CBO estimated an output multiplier for UI benefits ranging between 0.4 and 2.1—with the larger estimate being more relevant when monetary policy is at the zero lower bound.

How large a fiscal multiplier is needed to rationalize the gap between the micro and macro effects of the UI extensions? For this back-of-the-envelope exercise, we assume that the gap between the micro and macro estimates, $\Delta E_{MACRO} - \Delta E_{MICRO}$, arises solely due to aggregate demand effects. Since the multiplier is the ratio of total dollars created to total dollars spent, we first convert the employment effect of increasing

[40] As we noted in the introduction, Johnston and Mas (2015) provide a case study of Missouri where there was a sudden reduction in benefits, and find a much larger micro-level response than most of the literature. Besides providing labor supply based estimates, they also provide synthetic control and difference-in-difference estimates for aggregate employment effects from the benefit reduction. These macro estimates are similarly sized as their micro estimates, and are much larger than the macro effects that we find in this paper. Therefore, the size of the estimates from Johnston and Mas (2015) seem less about the micro versus macro effects than about the Missouri case study. Nonetheless, here we include the implied β_{MICRO} estimates from Johnston et al. study since those are specifically based on the labor supply response to the policy change.

[41] The exception is when we take the lower bound of the β_{MICRO} estimate (0.1) and the lower bound of the β_{MACRO} estimate (-0.1).

UI from 26 to 99 weeks into an impact on overall income, and then divide by UI expenditures. Our estimate of the change in total income is the product of the employment change (rescaling the percentage point change by $\frac{1}{100}$) and the ratio of output to employment $\left(\frac{Y}{E} = \$108,000\right)$.[42] National EB and EUC transfer payments between November 2008 and December 2013 averaged \$49.3 billion annually, and during this time period the average number of weeks of UI available was 74.4. In order to obtain an estimate of UI expenditures corresponding to an increase from 26 to 99 weeks, we scale the actual expenditure by $\frac{99-26}{74.4-26}$ $(\Delta B = \$49.3 \times 10^9 \times \frac{73}{48.4})$.[43] Dividing the estimated change in total income by the estimated UI expenditure gives our estimate of the fiscal multiplier, m_f.

$$
\begin{aligned}
m_f &= \frac{Y}{E} \times \frac{\Delta E_{MACRO} - \Delta E_{MICRO}}{100} \times \frac{1}{\Delta B} \\
&= \$108,000 \times \frac{(\beta_{MACRO} \times 253 + \beta_{MICRO} \times 134) \times 10^6}{100} \times \frac{1}{\$49.3 \times 10^9 \times \frac{73}{48.4}} \\
&= 3.7 \times \beta_{MACRO} + 1.9 \times \beta_{MICRO}
\end{aligned}
$$

In **Table 7**, we find that the implied fiscal multipliers using the first four micro estimates range between -0.2 and 2.3, centered around 1. However, when we use pre-Great Recession micro estimates, our implied multipliers are substantially larger, and range between 4.2 and 5.3. Finally, if we use the micro estimates from Johnston and Mas (2015), our implied fiscal multipliers are extremely large, exceeding 8. Since our macro effects are small, modest micro effects suggest a modest multiplier. However, large negative micro effects require a counterbalancing large fiscal multiplier to rationalize the small macro effect.

A caveat about our estimates is that our employment effects are estimated locally, and may differ from national multipliers for a number of reasons. First, a substantial fraction of the increased spending from UI extension is likely on tradable goods, much of which is produced outside of the local area. We are not capturing these demand leakages in our local analysis. Since a US county is substantially more open than the US as a whole, our local multiplier estimates are, *ceteris paribus*, likely to be smaller (and possibly substantially so) than national multipliers. Second, the multipliers estimated here are "transfer multipliers" as they are financed by transfers to the state from other states as opposed to through taxes or borrowing (Farhi and Werning (forthcoming)). Therefore, the transfer multiplier may reflect a wealth effect which would not be present at the national level when the spending is deficit-financed, making the local multiplier

[42] GDP per worker data from 2012 is from the World Bank: http://data.worldbank.org/indicator/SL.GDP.PCAP.EM.KD?locations=US. Note that this implicitly assumes that jobs created from the fiscal stimulus have mean productivity. Chodorow-Reich (2016) provides evidence supporting the validity of this approximation. Assuming that capital is fixed, but hours and employment adjust, he derives the following relationship between change in output and change in (headcount) employment: $\Delta Y \approx \theta \times (1+\chi) \times \frac{Y}{E} \times \Delta E$, where θ is labor's share, while χ is the elasticity of hours with respect to (headcount) employment. Given his estimates of $\chi = 0.5$ and $\theta = 0.7$, the constant-capital and hours adjustment channels cancel each other out, implying $\Delta Y \approx \frac{Y}{E} \times \Delta E$. He also validates the rough approximation using multipliers estimated from ARRA stimulus on state level employment and output. Similarly, Nakamura and Steinsson (2014) report both output and employment multipliers using defense spending shocks, and the magnitudes of both are are consistent with this approximation.

[43] We obtain the data for payments made through the EB and EUC programs from http://oui.doleta.gov/unemploy/euc.asp.

larger than the national multiplier. Nonetheless, Farhi and Werning (forthcoming) and Chodorow-Reich (2016) point out that as an empirical matter, externally funded transfer multipliers may provide a rough lower bound for the national, deficit-financed multipliers during liquidity traps. This is especially true when the transfer is not highly persistent, which was indeed the case for UI extensions. Overall, our estimates imply that a moderately sized, positive multiplier can rationalize the difference between the macro and the micro effects of UI, suggesting that the optimal benefit duration is likely to be countercyclical. This implication is consistent with the arguments in Landais et al. (2015) and Kroft and Notowidigdo (2016).

There are two other potential explanations for the gap between the micro and macro effects that come from recent work in search theory. The standard Pissarides (2000) search and matching model predicts that a higher benefit duration raises the negotiated wage, thereby reducing vacancies and employment through the job-creation effect (HKMM). However, if jobs are rationed, then a decrease in labor supply by some unemployed individuals from a more generous UI policy will tend to increase the job-finding probability of other unemployed workers, which can increase labor market tightness through the rat race effect (e.g., Michaillat (2012), Landais et al. (2015)). From the search-and-matching perspective, the net effect on employment will be a combination of the direct labor supply effect, the job creation effect, and the rat race effect. The positive macro effect, ΔE_{MACRO}, cannot be explained by the labor supply (which is negative), the job creation (which is negative) and rat race effects alone (which is positive but merely attenuates the negativity of the former two effects). As a result, it is indicative of at least some positive stimulative effect that may offset the negative effects from job creation and labor supply effects. However, the imprecision of the gap between the micro and the macro estimates suggests caution against interpreting this evidence too strongly.[44] Better distinguishing the search and aggregate demand channels remains an important area for future research.

7 Conclusion

Despite a large literature that has evaluated the labor supply effects of unemployment insurance, the overall impact of the policy on aggregate employment is a relatively new and understudied area of research. Yet, this is an important question from a public policy perspective. If there are sizable negative effects of UI employment via labor supply, but these are counteracted by positive aggregate demand effects, the overall employment effects can be more positive than what is implied by the labor supply estimates—making the policy more effective. Conversely, if the labor supply effects are small, but higher reservation wages fuels

[44]Many of our implied employment effects are not statistically distinguishable at a 95% level of confidence from the micro effects. However, for our PTT-Trimmed full sample specification, the 90% confidence interval does not contain the employment impacts implied by Rothstein's upper bound or any of the more negative micro estimates in **Table 7**.

lower hiring and hence a higher unemployment rate, the policy can be less attractive than may initially appear.

In this paper, we add to the small but growing literature on the impact of UI on overall employment. We utilize variation across counties which straddle state borders where the states differ in their UI duration during the Great Recession. We find that this strategy substantially reduces likely bias from endogeneity that would plague a two-way fixed effects model assuming parallel trends across counties (or states) receiving differential treatment. To account for remaining endogeneity, we utilize a variety of strategies including refining our sample and focusing on variation driven by the national policy changes created by the introduction of differential EUC in 2008 as well as the expiration of the EUC program at the end of 2013.

Whether we use all policy variations, or whether we use variation induced solely by national level policy changes, most of our estimates are quite small in magnitude. Our OLS results using a refined border county pair design suggest the employment to population ratio rose by a statistically insignificant 0.21 due to the 73 week increase in benefits. The IV results that use the national policy variation from 2008 expansion and 2014 expiration of EUC suggests the EPOP ratio changed by -0.07. While the 95% confidence intervals for the OLS estimate rules out change in EPOP more negative than -0.32, the confidence bounds for the IV rule out changes more negative than -1.31. Across a variety of specifications and samples, our preferred point estimates suggest that the extension of unemployment insurance duration from 26 to 99 weeks during the Great Recession led to a change in EPOP between -0.18 and 0.43 percentage points. Finally, our dynamic specifications do not indicate any policy anticipation effects.

Overall, our findings are similar to recent estimates by Chodorow-Reich and Karabarbounis (2016) and Coglianese (2015), who use policy variation that is quite different from what we use in this paper. At the same time, our estimates and conclusions are quite different from those reached by HMM and HKMM, even though they also use a border county pair based strategy. As we show in our **Online Appendix A**, the differences are in large part due to their use of (model-based) LAUS data, as well as auxiliary parametric assumptions used by authors of the two papers which we do not find to be warranted by the data.

The small macro employment effects of UI found in this paper are consistent with small negative effects on labor supply typically (though not always) found in the existing literature, together with moderately sized, positive effect on aggregate demand in the local economy. Future research should better disaggregate the macro effect into its constituent components: labor supply, demand multiplier, rat race and job creation effects. Nonetheless, our results suggest that the overall employment impact of the sizable UI extensions during the Great Recession was likely modest. At worst they led to a small reduction in aggregate employment, and at best they slightly boosted employment in the local economy.

References

Bai, Jushan. 2009. "Panel data models with interactive fixed effects," *Econometrica*, 77(4): 1229–1279.

Chodorow-Reich, Gabriel. 2016. "Geographic Cross-Sectional Fiscal Multipliers: What Have We Learned?" mimeo, Harvard University.

Chodorow-Reich, Gabriel and Loukas Karabarbounis. 2016. "The Limited Macroeconomic Effects of Unemployment Benefit Extensions," NBER Working Paper 22163, National Bureau of Economic Research, Inc.

Coglianese, John. 2015. "Do Unemployment Insurance Extensions Reduce Employment?" mimeo, Harvard University.

Congressional Budget Office. 2012. "Unemployment Insurance in the Wake of the Recent Recession,"Technical report.

Daly, Mary C., Bart Hobijn, Ayşegül Şahin, and Robert G. Valletta. 2012. "A Search and Matching Approach to Labor Markets: Did the Natural Rate of Unemployment Rise?" *Journal of Economic Perspectives*, 26(3): 3–26.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum wage effects across state borders: Estimates using contiguous counties," *The Review of Economics and Statistics*, 92(4): 945–964.

——— 2016. "Minimum Wage Shocks, Employment Flows, and Labor Market Frictions," *Journal of Labor Economics*, 34(3): 663–704.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin. 2010. "The Labor Market in the Great Recession," *Brookings Papers on Economic Activity*: 1–48.

Farber, Henry S., Jesse Rothstein, and Robert G. Valletta. 2015. "The Effect of Extended Unemployment Insurance Benefits: Evidence from the 2012–2013 Phase-Out," *The American Economic Review*, 105(5): 171–176.

Farber, Henry S. and Robert G. Valletta. 2015. "Do extended unemployment benefits lengthen unemployment spells? Evidence from recent cycles in the US labor market," *Journal of Human Resources*, 50(4): 873–909.

Farhi, Emmanuel and Iván Werning. forthcoming. "Fiscal Multipliers: Liquidity Traps and Currency Unions," *Handbook of Macroeconomics*: 2418–2491.

Hagedorn, Marcus, Fatih Karahan, Iourii Manovskii, and Kurt Mitman. 2015. “Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects.”

Hagedorn, Marcus, Iourii Manovskii, and Kurt Mitman. 2016. “Interpreting Recent Quasi-Experimental Evidence on the Effects of Unemployment Benefit Extensions,” NBER Working Paper 22280, National Bureau of Economic Research, Inc.

Johnston, Andrew and Alexandre Mas. 2015. “Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-Level Response to a Benefit Cut,” Unpublished manuscript, Princeton University.

Kroft, Kory and Matthew J Notowidigdo. 2016. “Should unemployment insurance vary with the unemployment rate? Theory and evidence,” *The Review of Economic Studies*: 1092–1124.

Lalive, Rafael, Camille Landais, and Josef Zweimüller. 2015. “Market externalities of large unemployment insurance extension programs,” *The American Economic Review*, 105(12): 3564–3596.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2015. “A Macroeconomic Theory of Optimal Unemployment Insurance,” mimeo.

Marinescu, Ioana. 2015. “The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board,” Unpublished manuscript, University of Chicago.

Michaillat, Pascal. 2012. “Do matching frictions explain unemployment? Not in bad times,” *The American Economic Review*, 102(4): 1721–1750.

Mitman, Kurt and Stanislav Rabinovich. 2014. “Do Unemployment Benefit Extensions Explain the Emergence of Jobless Recoveries?” Unpublished manuscript.

Nakamura, Emi and Jon Steinsson. 2014. “Fiscal stimulus in a monetary union: Evidence from US regions,” *The American Economic Review*, 104(3): 753–792.

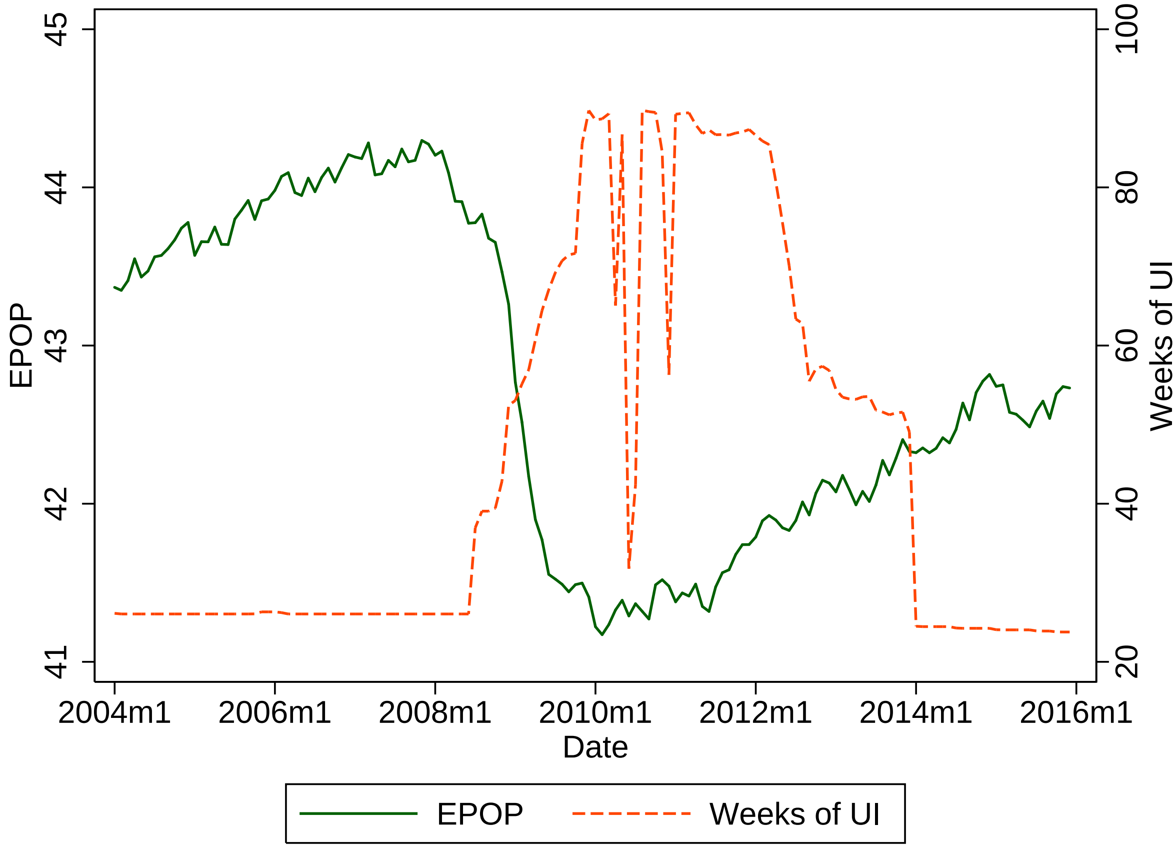
Pissarides, Christopher A. 2000. *Equilibrium unemployment theory*: MIT press.

Rothstein, Jesse. 2011. “Unemployment insurance and job search in the Great Recession,” *Brookings Papers on Economic Activity*: 143–214.

Sahm, Claudia R, Matthew D Shapiro, and Joel Slemrod. 2012. “Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?” *American Economic Journal: Economic Policy*, 4(3): 216–250.

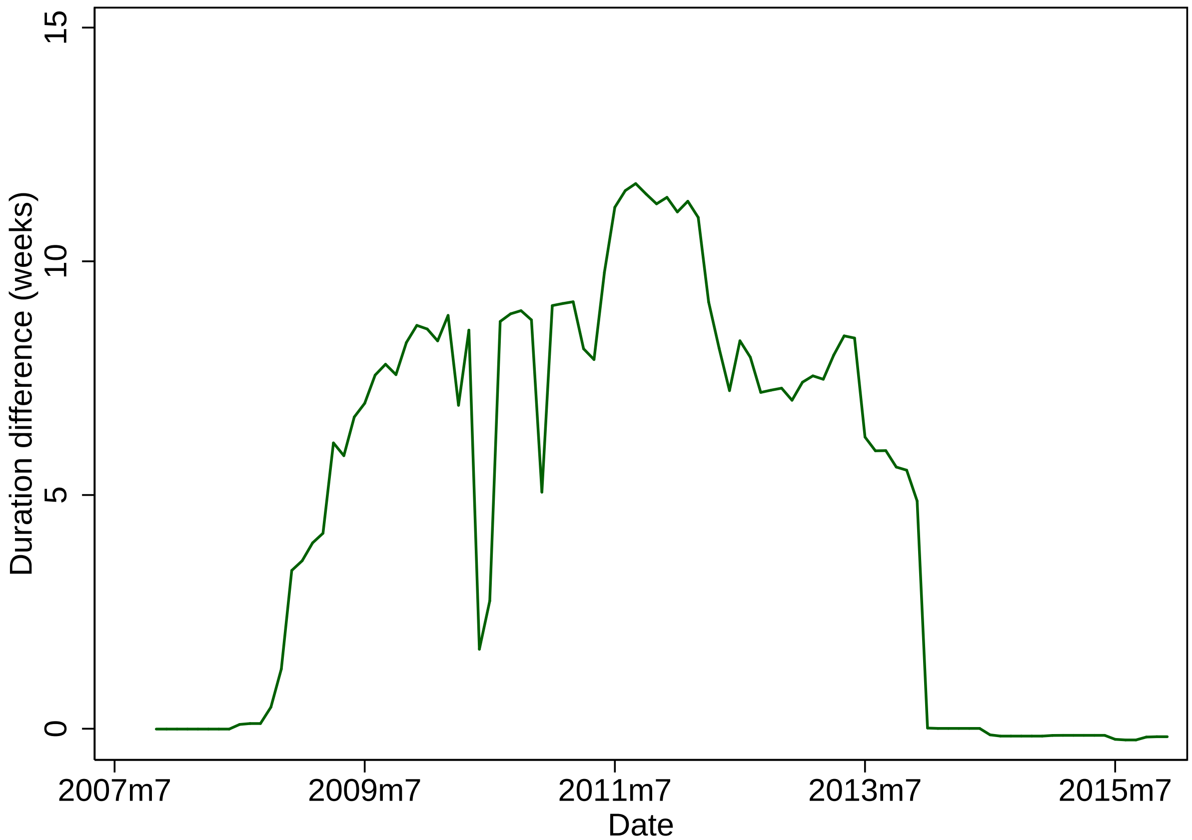
Summers, Lawrence H. 2010. "The Economic Case for Extending Unemployment Insurance," The White House Blog, July 14.

Figure 1: Evolution over time: national QCEW-based EPOP ratio and UI benefit duration



Notes: EPOP is the seasonally-adjusted ratio of employment (from the QCEW) to population age 15+. Weeks of UI represents the maximum number of weeks of UI compensation available. In this figure, both EPOP and weeks of benefits are calculated via an unweighted average of counties.

Figure 2: Difference in UI benefit duration between high-treatment and low-treatment counties across state borders



Notes: For each county pair, we compute the difference between maximum duration in the high-duration county and in the low-duration county. We plot the average difference across all county pairs. "High" and "low" status is determined by comparing the difference between average duration from 2008m11-2013m12 and average duration from 2007m11-2008m10 and 2014m1-2014m12. The counties in the 30 pairs where this difference is identical are assigned arbitrarily to the "high" and "low" sets.

Figure 3: Reduction in UI benefit duration from the December 2013 expiration of EUC

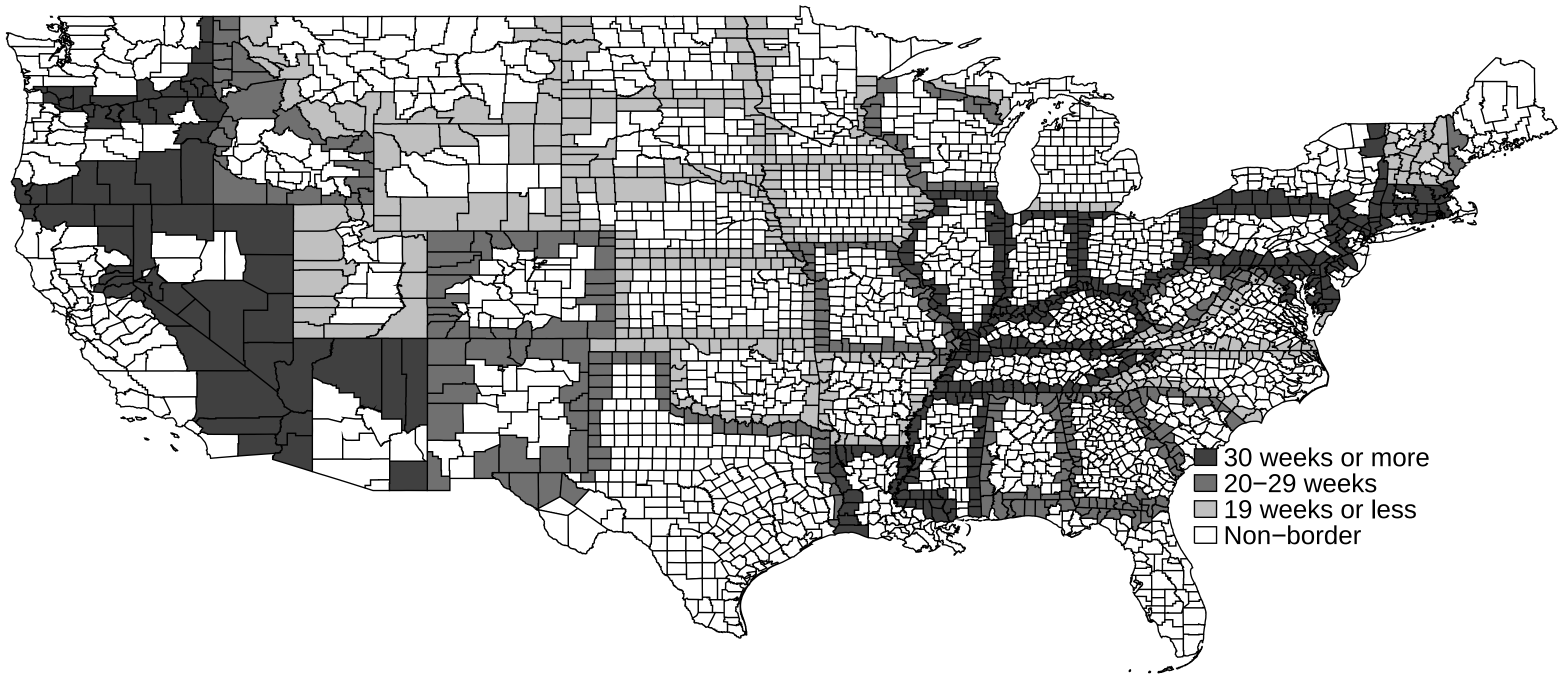
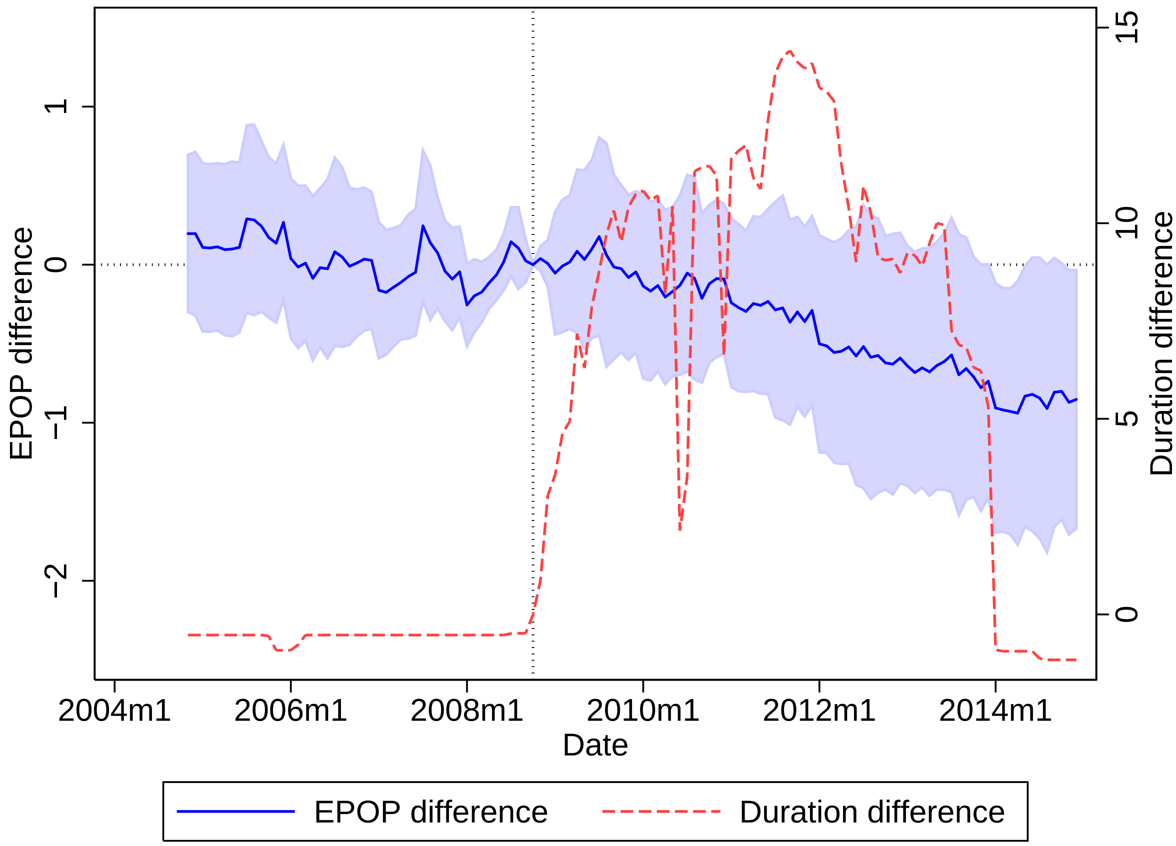
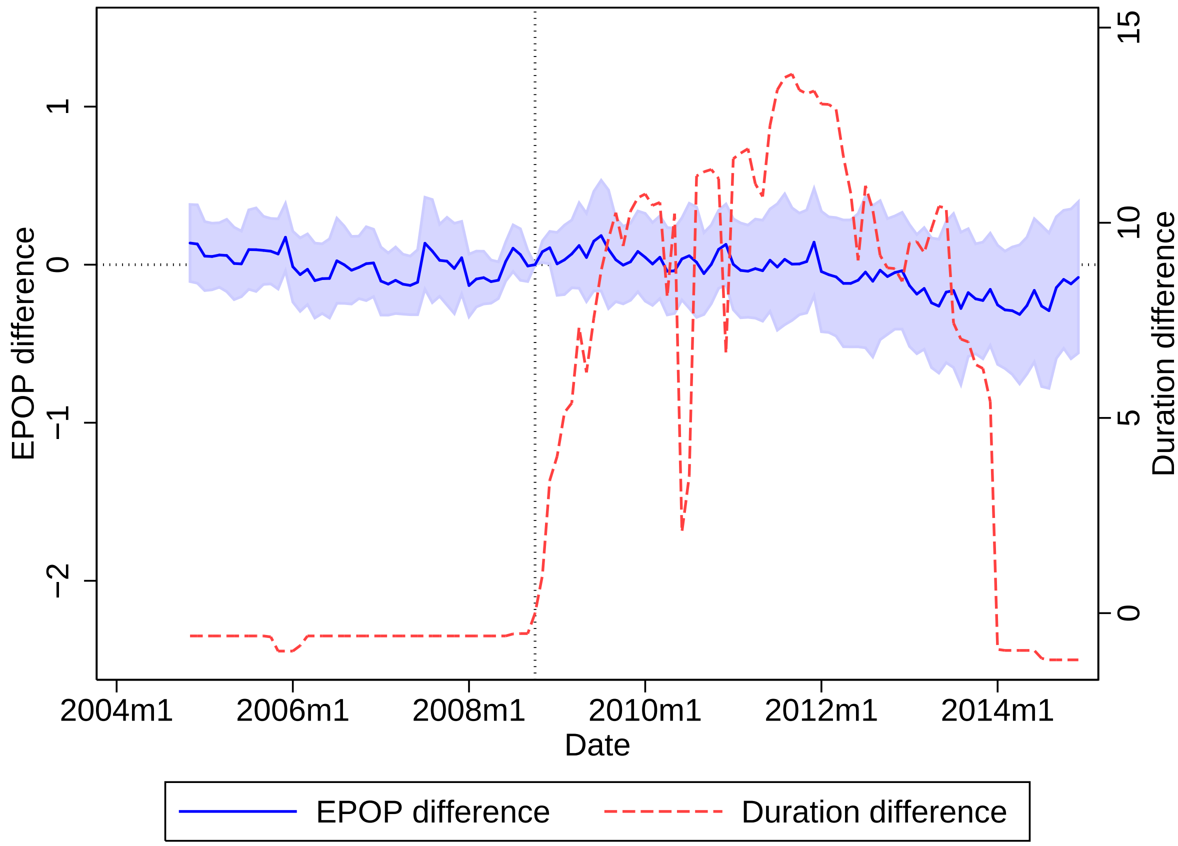


Figure 4: Evolution of EPOP and UI benefit duration differentials by average treatment intensity: baseline border county pair sample



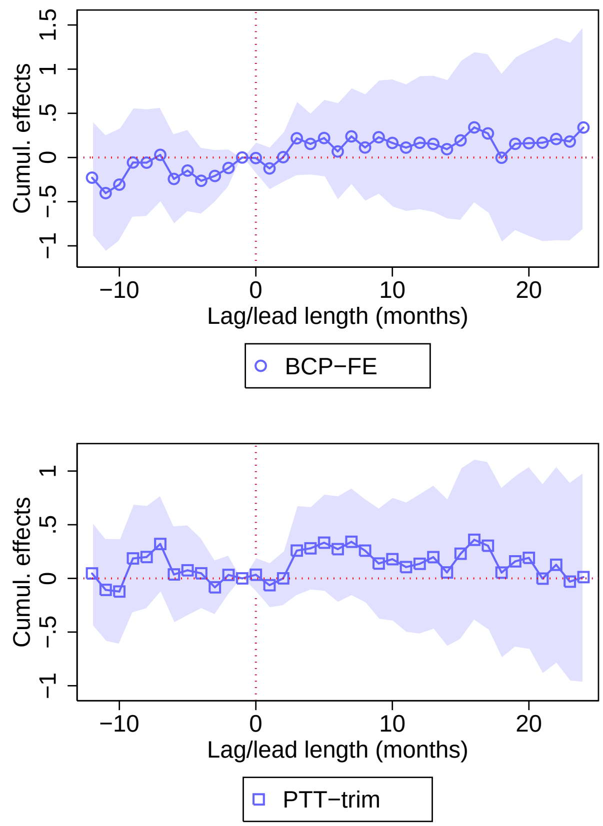
Notes: This figure plots (solid line, left axis) the set of β_s coefficients from the following regression: $E_{cpt} = \sum_{s=\tau_A}^{\tau_B} \beta_s treat_c \mathbb{1}\{t = s\} + \lambda_c + \nu_{pt} + \epsilon_{cpt}$. E_{cpt} is the seasonally-adjusted ratio of total employment to population age 15+, scaled in percentage points. The average treatment intensity, $treat_c$, is a time-invariant, continuous measure defined as the average duration during the treatment period (2008m11-2013m12), minus average duration from the 12 months prior (2007m11-2008m10), divided by 10. The shaded region corresponds to the 95% confidence interval, robust to two-way clustering at the state and state-pair level. The dotted line (right axis) reflects the analogous coefficients with D_{ct} as the dependent variable, where D_{ct} is weeks of benefits. The month 2008m10, the last month prior to the first introduction of differential EUC, is marked with a dotted vertical line. The sample includes 1,161 county pairs.

Figure 5: Evolution of EPOP and UI benefit duration differentials by average treatment intensity: PTT-trimmed border county pair sample



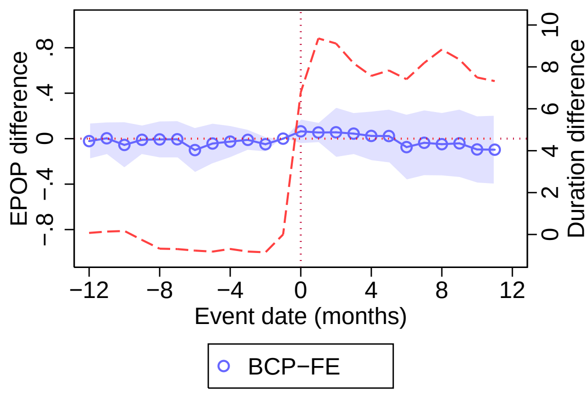
Notes: This figure plots (solid line, left axis) the set of β_s coefficients from the following regression estimated over the set of border county pairs in the PTT-trimmed sample: $E_{cpt} = \sum_{s=\tau_A}^{\tau_B} \beta_s treat_c \mathbb{1}\{t = s\} + \lambda_c + \nu_{pt} + \epsilon_{cpt}$. E_{cpt} is the seasonally-adjusted ratio of total employment to population age 15+, scaled in percentage points. The average treatment intensity, $treat_c$, is a time-invariant, continuous measure defined as the average duration during the treatment period (2008m11-2013m12), minus average duration from the 12 months prior (2007m11-2008m10), divided by 10. The shaded region corresponds to the 95% confidence interval, robust to two-way clustering at the state and state-pair level. The dotted line (right axis) reflects the analogous coefficients with D_{ct} as the dependent variable, where D_{ct} is weeks of benefits. The month 2008m10, the last month prior to the first introduction of differential EUC, is marked with a dotted vertical line. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. The sample includes 870 county pairs.

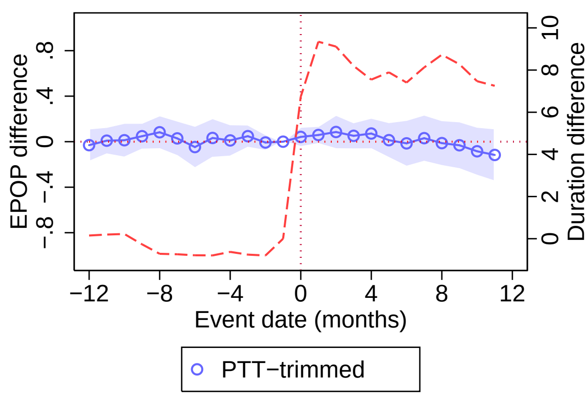
Figure 6: Cumulative response of EPOP from distributed lags specification: OLS in first-differences



Notes: This figure reports the monthly cumulative response of EPOP from a 73 week increase in maximum UI benefit duration, centered around event date -1 whose cumulative response is defined as zero. The model is estimated on the full sample (2007m11-2014m12), using all border county pairs (BCPs) (hollow circles) and the subset of BCPs in the PTT-trimmed sample (hollow squares), where all independent variables are divided by 73. The dependent variable is the first-differenced seasonally adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression includes 24 lags and 11 leads in first-differenced benefit duration, and is estimated using EPOP data from 2007m11-2014m12 (and thus duration data from 2005m11-2015m11). Lags are to the right of zero; leads are to the left of zero. The zeroth cumulative response is equal to the estimated coefficient on contemporaneous benefit duration. The j^{th} cumulative lag is equal to the estimated coefficient on contemporaneous duration plus the sum of the estimated coefficient on the 1st through jth lag term. The j^{th} cumulative lead is equal to -1 times the sum of the estimated coefficients on the first through the $j-1^{th}$ lead terms. The shaded region corresponds to the 95% confidence interval, robust to two-way clustering at the state and state-pair level.

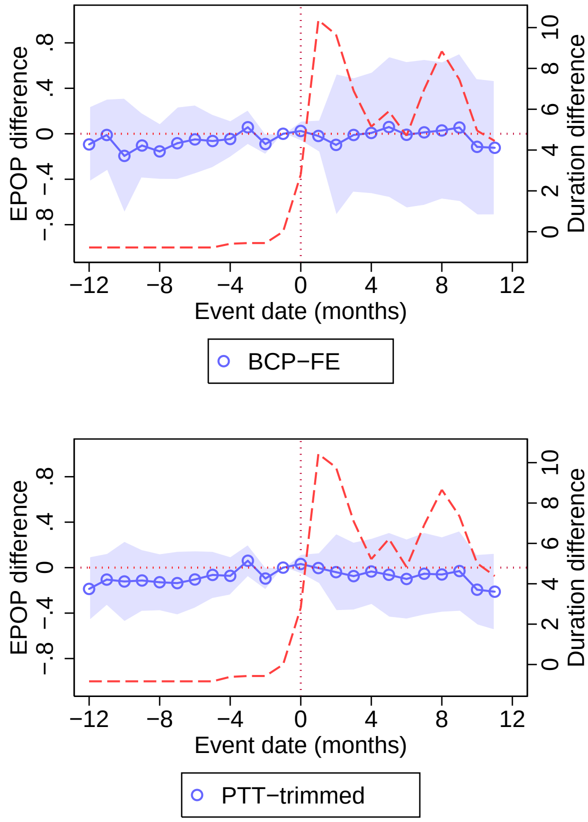
Figure 7: Evolution of EPOP difference and UI benefit duration difference across state borders: Pooled 2008 expansion and 2014 expiration of EUC





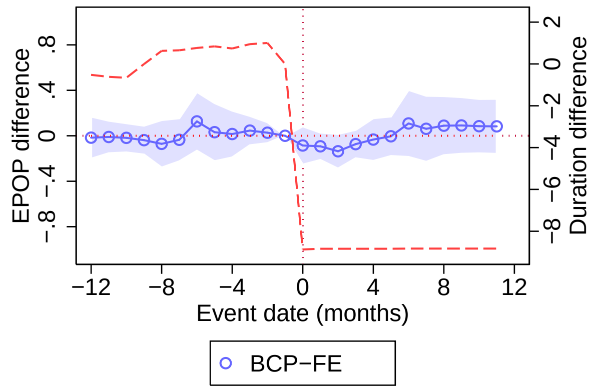
Notes: This figure reports the monthly cumulative response of EPOP (left axis, hollow circles) from the pooled 2008 and 2014 samples, centered around event date -1 whose cumulative response is defined as zero. The dependent variable is the first-differenced seasonally adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression includes 11 lags and 12 leads in first-differenced benefit duration: for the 2008 sample, the duration variable is equal to the increase in weeks of UI duration immediately upon the implementation of UCEA, divided by 10; for the 2014 sample, the duration variable is defined as -1 times the weeks of UI duration lost as a result of EUC expiration, divided by 10. The dashed line (right axis) reports the monthly cumulative response of benefit duration around the event; the regression is identical to the EPOP specification except that the dependent variable is the first-differenced benefit duration in weeks. The upper panel reports the results from the baseline BCP-FE sample consisting of all border county pairs; the lower panel reports results using the PTT-trimmed sample, which drops the quartile of county pairs with the highest differential in pre-treatment linear trends between November 2004 and October 2008. Event date zero is marked with a dotted vertical line; this corresponds to November 2008 for the 2008 sample and January 2014 for the 2014 sample.

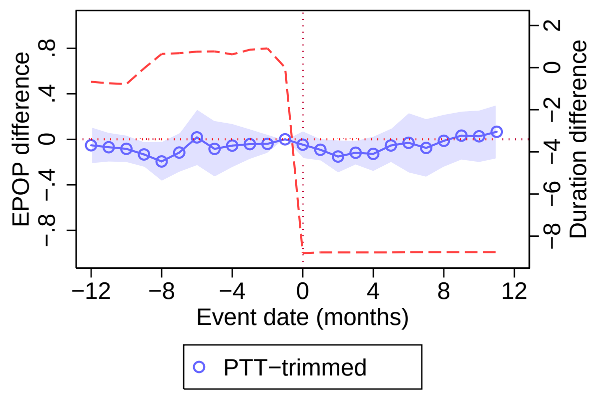
Figure 8: Evolution of EPOP difference and UI benefit duration difference across state borders: 2008 expansion of EUC



Notes: This figure reports the monthly cumulative response of EPOP (left axis, hollow circles) from the November 2008 EUC expansion, centered around event date -1 whose cumulative response is defined as zero. The dependent variable is the first-differenced seasonally adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression includes 11 lags and 12 leads in first-differenced benefit duration, where this duration variable is equal to the increase in weeks of UI duration immediately upon the implementation of UCEA, divided by 10. The dashed line (right axis) reports the monthly cumulative response of benefit duration around the event; the regression is identical to the EPOP specification except that the dependent variable is the first-differenced benefit duration in weeks. The upper panel reports the results from the baseline BCP-FE sample consisting of all border county pairs; the lower panel reports results using the PTT-trimmed sample, which drops the quartile of county pairs with the highest differential in pre-treatment linear trends between November 2004 and October 2008. Event date zero is marked with a dotted vertical line, and corresponds to November 2008.

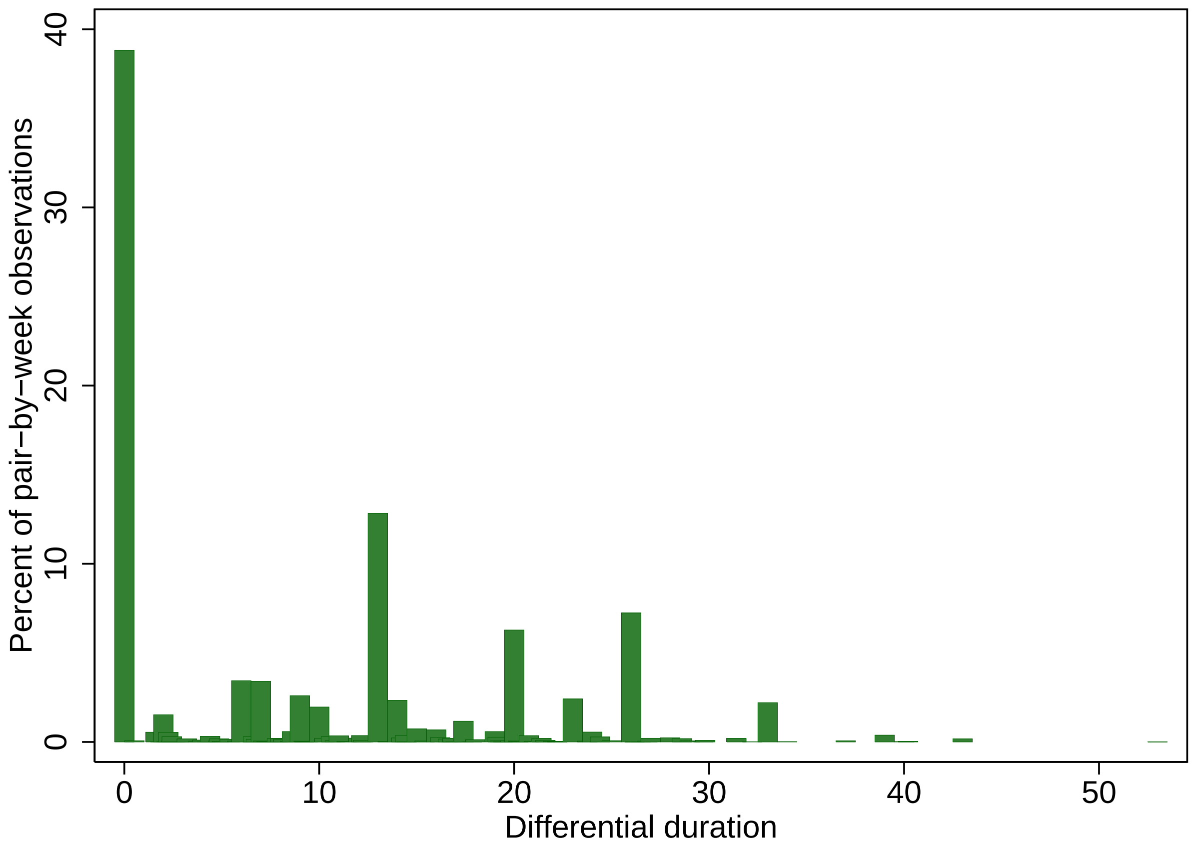
Figure 9: Evolution of EPOP difference and UI benefit duration difference across state borders: 2014 expiration of EUC





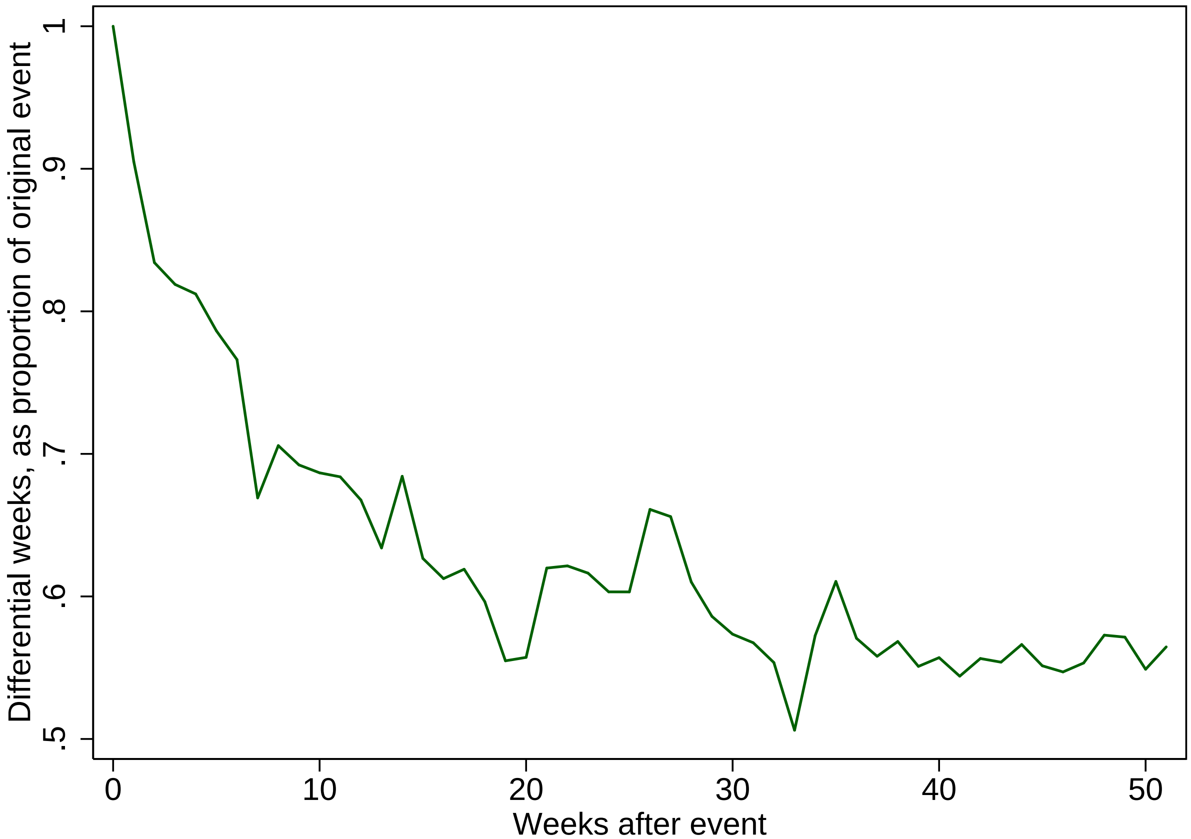
Notes: This figure reports the monthly cumulative response of EPOP (left axis, hollow circles) from the EUC expiration at the end of 2013, centered around event date -1 whose cumulative response is defined as zero. The dependent variable is the first-differenced seasonally adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression includes 11 lags and 12 leads in first-differenced benefit duration, where this duration variable is defined as change in weeks available as an immediate result of EUC expiration, divided by 10. The dashed line (right axis) reports the monthly cumulative response of benefit duration around the event; the regression is identical to the EPOP specification except that the dependent variable is the first-differenced benefit duration in weeks. The upper panel reports the results from the baseline BCP-FE sample consisting of all border county pairs; the lower panel reports results using the PTT-trimmed sample, which drops the quartile of county pairs with the highest differential in pre-treatment linear trends between November 2004 and October 2008. Event date zero is marked with a dotted vertical line, and corresponds to January 2014.

Figure 10: Distribution of differences in UI benefit duration across border county pairs



Notes: This figure plots the distribution of duration differences across border county pairs, with each observation at the pair-by-(calendar)-week level. The sample is restricted to weeks between November 23, 2008, and December 22, 2013.

Figure 11: Persistence of differential change in UI benefit duration across border county pairs



Notes: This figure plots the persistence of all changes in relative duration in the full sample. In particular, the data is organized at the pair (p), event (s), event-week (τ) level, where an event is any change in the duration difference across a county pair. The dependent variable $y_{ps\tau}$ is the difference in duration across the county pair, minus that same difference immediately prior to the event. This dependent variable is regressed on the size of the initial event interacted with 52 dummies for the 52 event-weeks τ immediately following the event. This figure plots those coefficients. See text for details.

Table 1: Pre-existing employment trends prior to November 2008 UI benefit expansion

	(1) All counties	(2) Border counties	(3) Border counties	(4) PTT-trimmed
Treatment X Date	-0.780***	-0.976***	-0.241	-0.110
	(0.244)	(0.206)	(0.286)	(0.110)
Observations	148896	111456	111456	83520
County fixed effects	X	X	X	X
Pair-period fixed effects			X	X

Notes: In columns 1 and 2, each cell reports the coefficient on $treat_c \times t$ from a regression of the following form: $E_{ct} = \alpha \times treat_c \times t + \lambda_c + \theta_t + \epsilon_{ct}$. In columns 3 and 4, each cell reports the coefficient on $treat_c \times t$ from a regression of the following form: $E_{cpt} = \alpha \times treat_c \times t + \lambda_c + \nu_{pt} + \epsilon_{cpt}$. In all columns, the dependent variable is the seasonally-adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression is estimated over the period 2004m11-2008m10 and t is the date divided by 48 (representing the 48 month period between the beginning and the end of this sample). The time-invariant variable $treat_c$ is the average treatment intensity for each county, defined as the average duration over the 2008m11-2013m12 period, minus average duration from the 12 months prior (2007m11-2008m10), divided by 10. In column 1, standard errors are clustered at the state level. In columns 2, 3, and 4, standard errors are clustered two-way at the state and state-pair level. Columns 4 report the estimates from the set of border county pairs in the PTT-trimmed sample. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 2: Summary statistics: High-treatment versus low-treatment counties in border county pair sample

	Baseline					**PTT-Trimmed**				
	High: Mean	Sd	Low: Mean	Sd	p-val	High: Mean	Sd	Low: Mean	Sd	p-val
EPOP (A)	44.227	17.113	44.741	15.269	0.679	42.983	14.641	44.075	13.721	0.285
Private EPOP (A)	34.773	16.168	35.132	14.567	0.762	33.723	13.652	34.804	13.342	0.243
LAUS unemp. rate (A)	5.127	1.674	4.864	1.843	0.135	5.241	1.585	4.916	1.794	0.023
Population age 15+ (A)	79,283	207,157	69,625	148,849	0.214	91,211	231,886	78,193	153,538	0.088
Share white (B)	0.811	0.182	0.811	0.177	0.998	0.815	0.180	0.818	0.176	0.838
Share black (B)	0.085	0.145	0.086	0.147	0.966	0.087	0.146	0.085	0.144	0.862
Share hispanic (B)	0.067	0.111	0.059	0.092	0.491	0.061	0.102	0.054	0.087	0.455
Share H.S. grad (B)	0.569	0.064	0.567	0.065	0.724	0.568	0.063	0.566	0.066	0.726
Share college (B)	0.179	0.078	0.189	0.086	0.010	0.182	0.080	0.193	0.088	0.000
Median h.h. income (B)	42,645	11,459	43,535	12,127	0.198	42,997	11,881	44,145	12,728	0.073
New mortgage debt p.c. (A)	3.456	3.226	3.674	3.039	0.423	3.556	2.961	3.836	3.090	0.251
Share in cities 50k+ (C)	0.190	0.331	0.196	0.331	0.759	0.203	0.338	0.222	0.348	0.267
Min. weeks of UI elig.	24.470	3.495	24.631	3.199	0.718	24.478	3.495	24.720	3.092	0.609
Max. weeks of UI elig.	96.105	6.674	86.996	13.320	0.000	96.452	6.212	87.755	12.787	0.000
Pairs w/ different avg treatment	1131		1131			849		849		
Pairs w/ identical avg treatment	30		30			21		21		

Notes: The first four columns report summary statistics in border counties in the estimation sample, separately for "high" and "low" treatment counties. A county's assignment to the "high" or "low" group is defined by its average treatment intensity relative to its counterpart within each pair. Average treatment intensity ($treat_c$) is a time-invariant, continuous measure defined as the average duration over the 2008m11-2013m12 period, minus average duration over the 2007m11-2008m10 and 2014m1-2014m12 periods. The 30 (baseline) or 20 (PTT-trimmed) border county pairs with identical treatment are dropped in this table. The fifth column reports the p-values from a test that the means for high counties and low counties are equal, robust to clustering two-way at the state and state-pair level. Columns 6-10 report analogous statistics for the subsample of border county pairs in the PTT-trimmed sample. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. If a border county appears in j county-pairs, then it appears j times for the purpose of creating the estimates in this table. (A) is from 2007 data, (B) is from the 2005-2009 ACS, and (C) is from the 2010 Census. High school graduates are those who have attained a high school degree but not a bachelor's degree. College graduates are those who have attained a bachelor's degree.

Table 3: Main Estimates: Effect of UI benefit duration on EPOP using OLS and IV specifications

	(1) BCP-FE	(2) PTT-Trimmed
Full sample		
OLS Estimate	0.430	0.213
	(0.466)	(0.270)
County pairs	1161	870
Observations	199692	149640
Pooled sample (IV)		
IV estimate	0.143	-0.069
	(0.964)	(0.635)
First stage coef.	0.847***	0.842***
	(0.052)	(0.051)
F stat.	[262.2]	[262.3]
County pairs	1161	870
Observations	108000	81120
2008 sample (IV)		
IV estimate	0.549	0.198
	(2.515)	(1.265)
First stage coef.	0.717***	0.726***
	(0.110)	(0.113)
F stat.	[41.3]	[40.3]
County pairs	1161	870
Observations	55728	41760
2014 sample (IV)		
IV estimate	-0.024	-0.182
	(0.562)	(0.521)
First stage coef.	0.915***	0.903***
	(0.046)	(0.043)
F stat.	[392.6]	[423.8]
County pairs	1089	820
Observations	52272	39360

Notes: Each panel reports two coefficients on D_{ct} from a regression of the form $E_{cpt} = \beta D_{ct} + \lambda_c + \nu_{pt} + \eta_{cpt}$. E_{cpt} is the seasonally-adjusted ratio of total employment to population age 15+, scaled in percentage points and D_{ct} is the potential weeks of UI benefits divided by 73. The second column restricts the sample to the PTT sample. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. Regressions in the first panel use OLS estimated over the 2007m11-2014m12 period. Regressions in the remainder of the table are estimated on subsamples using instrumental variables. The instrument z_{ct} is defined as follows. From 2007m11-2008m10, z_{ct} is equal to the duration available immediately prior to the implementation of UCEA; from 2008m11-2009m10, z_{ct} is equal to the duration available immediately *after* the implementation of UCEA. From 2013m1-2013m12, z_{ct} is equal to the duration available immediately prior to the expiration of EUC; from 2014m1-2014m12, z_{ct} is equal to the duration available immediately after EUC expiration, before any changes in regular benefits took effect. Estimates in the second panel pool the 2007m11-2009m10 and 2013m1-2014m12 samples and replace county fixed effects with county-by-subsample fixed effects. Estimates in the third panel use data from 2007m11-2009m10; estimates in the fourth panel use data from 2013m1-2014m12. In the IV specifications, first stage coefficients and standard errors are also reported. Standard errors are reported in parentheses and first stage F-statistics in square brackets. Standard errors are clustered two-way at the state and state-pair level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Robustness of the effects of UI benefit duration on EPOP: choice of sample period

	(1) BCP-FE	(2) PTT-Trimmed
2007m11-2014m12	0.430	0.213
	(0.466)	(0.270)
	N = 199692	N = 149640
2006m11-2014m12	0.142	0.175
	(0.451)	(0.322)
	N = 227556	N = 170520
2005m11-2014m12	-0.088	0.138
	(0.440)	(0.356)
	N = 255420	N = 191400
2004m11-2014m12	-0.330	0.064
	(0.452)	(0.378)
	N = 283284	N = 212280
County pairs	1161	870

Notes: Each cell reports the coefficient on D_{ct} from a regression of the form $E_{cpt} = \beta D_{ct} + \lambda_c + \nu_{pt} + \eta_{cpt}$. E_{cpt} is the seasonally-adjusted ratio of total employment to population age 15+, scaled in percentage points and D_{ct} is the potential weeks of UI benefits divided by 73. The second column restricts the sample to the PTT sample. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. The regression in each row is estimated over the sample-period indicated. The estimates in row 1 correspond to the estimates in the top panel of Table 3. Standard errors are clustered two-way at the state and state-pair level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 5: Robustness of the effects of UI benefit duration on EPOP: choice of cutoffs for trimming on match quality

	(1) Full	(2) Pooled	(3) 2008	(4) 2014
Baseline	0.430	0.143	0.549	-0.024
	(0.466)	(0.964)	(2.515)	(0.562)
	N = 199692	N = 108000	N = 55728	N = 52272
10th percentile	0.161	0.199	0.558	0.049
	(0.304)	(0.599)	(1.274)	(0.456)
	N = 179568	N = 97152	N = 50112	N = 47040
20th percentile	0.170	0.042	0.314	-0.074
	(0.276)	(0.612)	(1.241)	(0.507)
	N = 159616	N = 86544	N = 44544	N = 42000
25th percentile	0.213	-0.069	0.198	-0.182
	(0.270)	(0.635)	(1.265)	(0.521)
	N = 149640	N = 81120	N = 41760	N = 39360
30th percentile	0.221	-0.109	0.184	-0.232
	(0.272)	(0.629)	(1.216)	(0.549)
	N = 139664	N = 75648	N = 38976	N = 36672
40th percentile	0.329	0.085	0.929	-0.257
	(0.286)	(0.660)	(1.305)	(0.558)
	N = 119712	N = 64800	N = 33408	N = 31392
50th percentile	0.340	0.048	1.007	-0.328
	(0.302)	(0.719)	(1.435)	(0.601)
	N = 99760	N = 53952	N = 27840	N = 26112

Notes: Each cell reports the baseline coefficient from the full sample, pooled event sample, and 2008 and 2014 subsamples, estimated over a different subsample of border county pairs. The cells in row 1 correspond to the estimates in column 1 of Table 3. In the other rows, the sample of border county pairs (BCPs) is trimmed based on the magnitude of differences in pre-existing trends estimates from 2004m11-2008m10. We rank and then trim all BCPs according to the magnitude of differences in pre-treatment trends (PTT). In the second row, we drop the bottom 10 percent of BCPs with the largest differences in pre-existing trends, in the third row, we drop the bottom 20 percent, and so forth. The fourth row (the 25th percentile) corresponds to the estimates in column 2 of Table 3. Standard errors are clustered two-way at the state and state-pair level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Additional robustness checks on the effects of UI benefit duration on EPOP

	Full sample OLS		**Pooled sample IV**	
	(1) BCP-FE	(2) PTT-Trimmed	(3) BCP-FE	(4) PTT-Trimmed
1. Baseline	0.430	0.213	0.143	-0.069
	(0.466)	(0.270)	(0.964)	(0.635)
2. Private EPOP	0.268	0.058	0.205	0.090
	(0.489)	(0.269)	(1.012)	(0.631)
3. Correlation-trimmed	-0.145	-0.003	-0.120	-0.309
	(0.351)	(0.266)	(1.106)	(0.646)
4. ISLT	0.380	0.189	0.930	0.659
	(0.363)	(0.233)	(0.659)	(0.602)
5. PTT through 2007m10		0.326		0.166
		(0.295)		(0.615)
6. Eliminate lapse	0.543	0.258		
	(0.516)	(0.275)		
7. Quarterly data	0.453	0.239	0.205	0.026
	(0.507)	(0.300)	(0.918)	(0.617)
8. QWI EPOP (quarterly)	0.692	0.639^{*}	0.402	0.235
	(0.475)	(0.331)	(0.639)	(0.637)
9. Not seasonally adjusted	0.301	0.141		
	(0.481)	(0.277)		
10. Distance trimming	0.323	0.246	0.313	0.136
	(0.402)	(0.276)	(1.119)	(0.700)
11. Unbalanced panel	0.329	0.213	0.148	-0.069
	(0.469)	(0.270)	(0.947)	(0.635)
12. Exploit Δ reg. benefits			0.185	-0.039
			(0.950)	(0.621)
13. Drop NC	0.416	0.184	0.159	-0.050
	(0.550)	(0.315)	(0.994)	(0.655)
14. Keep NC			-0.147	-0.406
			(1.033)	(0.743)
15. NC: Alt. instrument			0.071	-0.081
			(0.623)	(0.425)
16. $\ln(EPOP)$	0.006	0.003	0.007	-0.004
	(0.008)	(0.007)	(0.017)	(0.011)
	[0.335]	[0.186]	[0.424]	[-0.232]
17. $\ln(emp)$	0.008	0.003	0.014	0.004
	(0.009)	(0.009)	(0.016)	(0.011)
	[0.429]	[0.166]	[0.802]	[0.209]

Notes: Each cell reports regressions analogous to those reported in Table 3 for the full sample with OLS or the pooled event samples (IV). The estimates in the 1st row correspond to the estimates in the top two panels of Table 3. The estimates in the 2nd row replace (total) EPOP with the ratio of private employment to population age 15+. In the 3rd row, we trim the set of border county pairs based on the level of correlation between county EPOP and state EPOP over the period 2004m11-2008m10 (see text for details). The 4th row controls for county-specific linear trends. The 5th row trims based on PTT estimated through 2007m10 instead of 2008m10. The 6th row recodes the periods in 2010 when EUC lapsed by assigning EUC values during these lapses as equal to their prior value. The 7th row uses quarterly data instead of monthly (and estimates over the 2007q4-2014q4 period). The 8th row uses EPOP derived from the QWI (at the quarterly level) instead of the QCEW. The 9th row uses seasonally-unadjusted data. The 10th row drops county-pairs whose population centroids are greater than 100km apart. The 11th row includes counties without full EPOP data for each month, which we drop by default. The 12th row uses a modified version of the instrument z_{ct} which exploits all changes in benefits, including changes in regular benefits, which occur at the end of December 2013. Rows 13-15 report estimates using alternative strategies for dealing with North Carolina (NC); by default, border county pairs (BCPs) with one neighbor in NC are kept in the full sample OLS and the 2008 subsample and dropped in the 2014 subsample. The 13th row completely drops all NC BCPs. The 14th row keeps all North Carolina BCPs. The 15th row keeps NC BCPs but redefines the instrument for NC counties (see text for details). The 16th and 17th row use $\ln(EPOP)$ and $\ln(employment)$, respectively, as dependent variables. The bracketed estimates in these two rows are the level-on-level equivalent, equal to $(\frac{99}{26}^{\hat{\beta}} - 1)\bar{E}$, where $\bar{E}$ is the mean EPOP level in the given sample. Cells which are not applicable in the given sample, or which provide estimates that are mechanically equal to the baseline estimates, are left blank. Standard errors are clustered two-way at the state and state-pair level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Rationalizing micro and macro employment effects of UI: demand side effects and implied fiscal multipliers

	β_{MICRO}	β_{MACRO}	ΔE_{MICRO}	ΔE_{MACRO}	m_f
Rothstein (2011), lower bound	0.1	-0.1	-0.1	-0.3	-0.2
	0.1	0.2	-0.1	0.5	0.9
Farber and Valletta (2015)	0.4	-0.1	-0.5	-0.3	0.4
	0.4	0.2	-0.5	0.5	1.5
Rothstein (2011), upper bound	0.5	-0.1	-0.7	-0.3	0.6
	0.5	0.2	-0.7	0.5	1.7
Daly et al. (2012)	0.8	-0.1	-1.1	-0.3	1.2
	0.8	0.2	-1.1	0.5	2.3
Elsby et al. (2010), upper bound	2.4	-0.1	-3.2	-0.3	4.2
	2.4	0.2	-3.2	0.5	5.3
Johnston and Mas (2016)	4.6	-0.1	-6.2	-0.3	8.4
	4.6	0.2	-6.2	0.5	9.5

Notes: The table displays estimates of implied fiscal multipliers, using a range of micro estimates from other studies (Column 1) and two macro estimates from this paper (Column 2). β_{MICRO} is an estimate of the change in the unemployment rate resulting from only the micro-level effect of a 73-week increase in maximum UI duration, while β_{MACRO} is a direct estimate of the aggregate change in EPOP. Columns (3) and (4) represent the resulting impact on employment (in millions of workers), and are calculated as $\Delta E_{\text{MICRO}} = \beta_{\text{MICRO}} \times L$ and $\Delta E_{\text{MACRO}} = \beta_{\text{MACRO}} \times P$, where P is the population and L is the labor force, expressed in millions. m_f is the implied fiscal multiplier, computed under the assumption that the entirety of the gap between the macro and micro employment effects is due to aggregate demand. m_f is computed according to the equation $m_f = \frac{Y}{E} \times \frac{\Delta E_{MACRO} - \Delta E_{MICRO}}{100} \times \frac{1}{\Delta B}$, where $\frac{Y}{E}$ is output per worker in 2012, and ΔB is annual EB and EUC expenditure, averaged over the period from November 2008 through December 2013 and scaled to correspond to a 73-week increase in duration.

Online Appendix A: Comparison with HKMM and HMM

The results in this paper are quite different than the results in Hagedorn, Karahan, Manovskii and Mitman (2015) (which studies the effect of UI from 2005 to 2012) and the results in Hagedorn, Manovskii and Mitman (2016) (which studies the effect of EUC expiration at the end of 2013). Similar to this paper, both HKMM and HMM use border county pairs for their estimation. However, there are differences in data, in econometric specification, and in sample definitions between our paper and these two studies. Some differences are minor, while others are quite important. In this online Appendix, we offer a reconciliation of these sets of results, and assess which of the differences in sample definition, data and specifications drive the differences in the estimates and conclusions reached by our respective papers.

Comparison to HKMM

In this section, we compare our OLS estimates from the baseline BCP sample to the baseline estimates of HKMM. The HKMM estimation equation is as follows, where data for a given pair p at time t has already been spatially differenced (after taking logs):

$$\ln(u_{pt}) - \beta(1 - s_t)\ln(u_{pt+1}) = \alpha * \ln(D_{pt}) + \lambda_p' F_t + \epsilon_{pt} \tag{A1}$$

Here, u_{pt} is the unemployment rate from LAUS,[45] β is the discount factor (equal to 0.99), s_t is the separation rate, D_{pt} is the same measure of maximum benefit lengths that we use, and $\lambda_p' F_t$ are interactive effects. Thus, the dependent variable is a quasi-forward difference (QFD) of the log of the unemployment rate. They then calculate the total effect of UI on unemployment by considering the steady state ($u_{pt} = u_{pt+1}$) impact of a persistent increase in D_{pt}. In the steady state, $\ln(u_p) = \frac{\alpha}{1-\beta(1-s)}\ln(D_p)$. Therefore, HKMM's headline claim comes from multiplying their main estimate by a factor $\frac{1}{1-\beta(1-s)}$, which is approximately equal to 10. They perform their estimation over the period 2005q1-2012q4.

Our full-sample BCP-FE estimation strategy is different from HKMM in five distinct ways. These differences are: (1.) we do not transform our dependent variable using quasi-forward-differencing, (2.) we use employment data from the QCEW rather than unemployment data from LAUS, (3.) we estimate the results using monthly data from 2007m11-2014m12, instead of quarterly data from 2005q1-2012q4, (4.) we control for differences across county pairs using a fixed effects model rather than the Bai (2009) interactive fixed effects model, and (5.) we use levels instead of logs.

[45] The LAUS data used by HKMM has been substantially revised since they accessed it. We have estimated the models using both the pre-revision version of the LAUS data used by HKMM and the more recent, revised version of the data. We have found both versions of the data give similar results in the HKMM specifications. We use pre-revision data throughout the discussion of HKMM.

Appendix Table A1 describes the impact of each of these five steps. Because different specifications have different dependent variables, and because the implied effect is not equal to the coefficient in some specifications, we standardize each specification into an implied effect of the 26-to-99 week expansion on EPOP.[46] We "translate" between implied effects on the unemployment rate and implied effects on EPOP by using the total peak-to-trough impact of the Great Recession. We measure this peak-to-trough impact using the unweighted average of counties in our border-pair sample. In particular, in this sample, EPOP fell from 44.3% to 41.2% and the unemployment rate increased from 4.8% to 9.7%. So, if one estimation suggests that the impact of the 26-to-99 week expansion was 3 percentage points of unemployment, we would convert that specification's estimate into an EPOP effect of $3 \times (\frac{41.2-44.3}{9.7-4.8}) \approx -1.9$ percentage points.

Appendix Table A1 analyzes one-off changes either starting from the HKMM specification (column 1), or moving to our specification (column 2). The first row begins with reporting the estimates: our replication of the HKMM estimates suggest that the UI benefit expansion from 26 to 99 weeks has an implied EPOP effect of -2.72, which is nearly 90% of the decrease in EPOP during the Great Recession within our sample. This corresponds to a coefficient estimate of 0.052, while HKMM report a very similar estimate of 0.049. We find that this estimate is statistically significant, as HKMM do. In contrast, the point estimates for the full sample BCP-FE estimates in this paper suggest that the decline in EPOP would have been about 10% *greater* without the UI expansions, though this is not distinguishable from zero.

The next five rows report the marginal impact of each of the five steps. In column 1, we show what happens when the step reported in the row is added starting with the HKMM specification. In column 2, we show what happens when this step is added to our specification. Finally, in column 3, we consider *all* possible transition paths between HKMM's estimates and our estimates, and report the average marginal contribution of each of the steps, across all of these transition paths.[47]

The key findings are as follows. Quasi-forward differencing, the use of the LAUS unemployment data as opposed to the QCEW employment data, and sample alignment are all consequential choices. In contrast, the use of interactive fixed effects as opposed to linear fixed effects and the use of logs versus levels are not consequential choices.

Column 1 shows that, starting from the HKMM estimate, switching from the LAUS unemployment rate, or getting rid of quasi-forward differencing, dramatically reduces the HKMM estimates in magnitude towards

[46] Importantly, we scale up the estimates in QFD specifications by $1/[1-\beta(1-s)]$, as HKMM do.

[47] We do not consider the step of switching from logs to levels in column 1, because the quasi-forward-differencing is motivated by theory which requires the data to be in logs. With quasi-forward-differenced data in levels, it is neither clear what we are measuring, nor what the total effect of UI on employment would be. For the same reason, we do not consider adding quasi-forward-differencing to our specification in column 2 (which is in levels). In addition, when calculating the averages in column 3, we discard transition paths that involve using quasi-forward-differenced data in levels. In the end, we estimate 24 models with all allowable combinations of the five sources of differences; we then take 60 paths (equal to 5! paths with 1/2 thrown out because eliminating quasi-forward differencing happens after the logs to levels conversion) between the HKMM and BDGK estimates, and calculate the contribution of each of these five factors averaged across these 60 paths.

zero. In particular, just switching from the LAUS unemployment rate to the QCEW EPOP (as shown in Row 4) changes the estimates to $-2.724 + 1.356 = -1.368$, suggesting the UI benefit expansion explained around 40% of the fall in EPOP rather than 90% as implied by HKMM's estimates. Similarly, removing quasi-forward differencing (Row 2) changes the estimates to $-2.724 + 2.688 = -0.036$ percentage points of EPOP. Column 2 shows that use of the LAUS unemployment rate also leads to a (mistaken) suggestion of job loss when we start from our specification, although the impact of this is more modest. Starting from our BCP-FE specification, when we use the LAUS unemployment rate as the outcome, the translated result suggests the UI benefit expansion led to a change in EPOP equal to $.430 - 1.133 = -0.703$, just under a quarter of the overall change during the Great Recession. When we average the incremental contribution of these two steps across all permissible paths going between the HKMM specification and ours (in column 3), we find that dropping quasi-forward differencing increases the estimates by around 1.32 percentage points of EPOP (about 40% of the change in unemployment rate during the Great Recession), while switching the outcome from LAUS unemployment rate to QCEW based EPOP increases the estimate by about 0.74 percentage points of EPOP.

Aligning our samples also has a meaningful impact. The HKMM sample of 2005q1-2012q4 starts and ends earlier than our sample of 2007m11-2014m12. As we showed in **Table 4**, while the baseline BCP-FE approach greatly reduces the pre-existing trend, it does not completely remove it. Use of an earlier start date, as well as an end date prior to the phase-out of differential UI benefits across state borders, can produce a more negative estimate in the presence of such trends. We find that use of this altered sample period leads to somewhat smaller magnitudes of estimates, reducing the impact of the policy by around 0.846, 1.461, and 0.863 percentage points of EPOP in columns 1, 2, and 3, respectively.

In contrast, the use of Bai (2009) interactive effects versus fixed effects, and use of logs versus levels, make fairly small contributions in explaining the difference between our two sets of estimates.

This analysis shows that (1) changing the sample period (and frequency) from HKMM's specification to ours, (2) eliminating quasi-forward-differencing, and (3) changing the dependent variable from the LAUS unemployment rate to QCEW EPOP all reduce the implied negative impact of UI on employment, by 0.74 to 1.25 percentage points of EPOP when averaged over all possible paths. We next discuss our justification for making the specification choices that we do.

Quasi-Forward Differencing

HKMM derive **Equation (A1)** by considering a search-and-matching framework where the rate of vacancy posting or firm job creation depends on a firm's expectation about future wages. Since unemployment insurance puts upward pressure on wages, an increase in benefits would reduce the expected profits of the firm and lead to a reduction in job creation. Because expectations about *future* benefit changes can affect employment *today*, HKMM make the point that an empirical approach that only relates current employment to current or past policy changes would be misspecified. In order to capture these anticipation effects, HKMM use a quasi-forward-differencing procedure. Their argument is as follows: the value of an employee to an employer is equal to the current-period flow profits, plus $\beta(1-s)$ times the expected value of the employee tomorrow (since the value of a vacant job is driven to zero by free entry). Therefore, HKMM argue, we can isolate the impact of UI on current-period flow profits by considering the quasi-forward difference of the unemployment rate (which they consider to be proportional to current period flow profits, in logs). The theory predicts that, in the case of an increase in generosity that was a surprise and immediately known to be persistent, firms would move from a low-unemployment steady state to a high-unemployment steady state, according to the equation $\Delta\ln(u_p) = \frac{\alpha}{1-\beta(1-s)}\Delta\ln(D_p)$.[48] As we noted above, this choice is quite important—removing forward differencing essentially erases the entirety of their effect even in their sample.

We are generally less favorable toward the use of quasi-forward differencing for several reasons. This model-driven approach relies on strong parametric assumptions—most notably that labor demand is well-characterized by the vacancy-posting problem captured in the model. Unfortunately this results in an empirical approach that is very sensitive to misspecification. For example, if an increase in UI generosity (D_{pt}) tends to be associated with a decrease in future unemployment (u_{pt+1}) in the data, then the estimated coefficient α will be positive. However, such a pattern could also be consistent with a Keynesian aggregate demand effect that operates with a small delay. That is, if an increase in benefits in one period leads to increased aggregate demand and lower unemployment in the next period, the HKMM strategy would find that UI *increased* the unemployment rate, when in fact the opposite occurred. Second, as a practical matter, the size of the final estimate is sensitive to assumptions in the model required for translating a flow result to a steady state effect, and in the exact magnitudes of separation and discount rates. Both the heavy dependence on a specific model and the inability to distinguish between alternative explanations make quasi-forward differencing an unattractive strategy from our perspective.

Instead, our preferred strategy is to capture the dynamics in a less model-driven and a more transparent manner using distributed lags. That specification directly estimates employment changes around benefit duration innovations, allowing us to assess possible pre-existing trends, anticipatory effects, and delayed or slow moving response within the window. As we discussed in Section 5.2, we find no evidence of significant anticipation effects in the 12 months prior to benefit changes. The lack of any anticipation effect raises

[48]Here α is the regression coefficient, β is the discount factor, s is the probability that the job ends, u is the unemployment rate, and D is the number of weeks of UI benefits.

questions about the value of quasi-forward differencing the outcome, especially given the drawbacks discussed above.

LAUS versus QCEW

HKMM predominantly use the LAUS employment data rather than the QCEW employment data to compute county level measures of employment.[49] However, the LAUS data is partly model-based. In particular, while the LAUS data uses actual movement to unemployment based upon UI claims, they do not observe those entering (or re-entering) the labor force. Therefore, the county level estimates for unemployment are based on state-level data on labor force entry and re-entry—something BLS states explicitly in their online manual (http://www.bls.gov/lau/laumthd.htm):

> "The second category, "new entrants and reentrants into the labor force," cannot be estimated directly from UI statistics, because unemployment for these persons is not immediately preceded by the period of employment required to receive UI benefits. In addition, there is no uniform source of new entrants and reentrants data for States available at the LMA [labor market area] level; the only existing source available is from the CPS at the State level. Separate estimates for new entrants and for reentrants are derived from econometric models based on current and historical state entrants data from the CPS. These model estimates are then allocated to all Labor Market Areas (LMAs) based on the age population distribution of each LMA. For new entrants, the area's proportion of 16-19 years population group to the State total of 16-19 years old population is used, and for reentrants, the handbook area's proportion of 20 years and older population to the State total of 20 years and older population is used."

The use of state-level information in estimating county-level unemployment rates is problematic for a border discontinuity design. The border county design attempts to purge reverse causation present at the state level by using more local comparisons. Use of state-level information raises the possibility of finding a (spurious) discontinuity in the measured unemployment rate across the state borders even when there is no such discontinuity in reality.

The QCEW data are based on administrative payroll records provided to the BLS by states, which protects against finding spurious discontinuities. Moreover, the QCEW data includes around 98% of all formal sector workers, making them very close to the true total employment counts in these counties. For

[49] They do report results using the log employment from the QCEW and QWI as a robustness check, in columns 3 and 4 of Table 5. The log employment result, -0.03, would imply that the 26-99 week expansion of UI caused a reduction of employment by 3.9%, which would translate to about 1.6 percentage points of EPOP. This is about 40% less than implied EPOP effect of HKMM's main result, consistent with the average marginal effects reported in **Appendix Table A1**. The log employment results from the QWI are modestly larger.

these reasons, we consider the QCEW to be the preferred data source for county-level employment. When the results using the QCEW and LAUS data differ non-trivially—which they do in this case—the QCEW findings are much more likely to be accurate.

Sample Alignment

HKMM's sample goes from 2005 through 2012 and uses quarterly data. By contrast, our main specification uses monthly data, starts in 2007m11, and goes through 2014m12. Using quarterly versus monthly data has virtually no impact. For our preferred specification, for example, changing to quarterly data increases the standard errors by a little more than 0.04 and increases the the mean estimate by 0.02 (see **Table 6**). Though that represents a 7% increase, since the baseline estimates are small to start with, the impact is quite small. However, switching the time period of estimation from 2005-2012 to 2007m11-2014m12 does make a difference. First of all, as we discussed in **Section 5.3**, the 2007m11-2014m12 sample exhibits a fairly symmetric rise and then fall in treatment intensity, orthogonalizing possible trends. Moving to the 2005-2012 sample makes this less so. As can be seen in **Figure 4**, the 2005-2012 period is largely a period of (1) increasing benefit duration and and (2) decreasing relative employment on the high-treatment side of the border. However, after 2012, the high-treatment side of the border starts to experience a relative decline in duration, while continuing its relative decline in employment. This is in part due to federal policy changes and in part due to differential changes in unemployment levels. Thus, it is not surprising that adding 2013 and 2014, and removing 2005 to 2007m10, has a noticeable positive impact on the UI duration impact upon employment.

Furthermore, we note that the choice of sample date matters little for the PTT-trimmed sample. **Table 4** shows that the OLS estimates in the BCP-FE specification fall from 0.41 to -0.35 when the sample is changed from 2007m11-2014m12 to 2004m11-2014m12. However, the OLS estimates in the PTT-trimmed sample fall only from 0.18 to 0.03. The IV estimates show a similar pattern, although the range is larger in both samples. This leads us to be confident that the large negative effects seen in full-sample specifications with earlier start dates (and/or end dates) reflect endogeneity from pre-existing trends. Furthermore, since the 2007m11-2014m12 sample window effectively orthogonalizes these trends with treatment, we believe that our sample window provides for more reliable estimates than other sample windows, including HKMM's 2005q1-2012q4.

HMM comparison

HMM find that the expiration of EUC at the end of 2013 increased employment, though the implied effect of UI generosity is smaller than that of HKMM. Whereas the latter suggests that approximately 80% of the increase in unemployment during the Great Recession can be explained by the increase in benefit generosity, applying the coefficient estimates of HMM to the 26-to-99 week expansion would imply that UI policy can explain about one third. Scaled another way, HMM finds that the employment effect of the *expiration* is on the same order as total employment gains during 2014. HMM estimate a variety of different empirical models, all of which are motivated by a desire to exploit variation in UI benefits solely coming from the EUC expiration, while at the same time incorporating information over a longer period to formulate a counterfactual for the county-level employment which would have occurred had EUC not expired. Broadly, these specifications can be broken into two groups, which we call the "interaction term" models and the "event study" models.[50] We discuss each of them in turn.

The following is equivalent to HMM's "benchmark" interaction term model, where e_{ct} is log employment, measured either in the QCEW or LAUS:[51]

$$e_{ct} = \kappa[\ln(D_{ct})\mathbb{1}(t \leq 2013q3)] + \alpha[\ln(D_{ct})\mathbb{1}(t \geq 2013q4)] + \mu_c + \nu_{pt} + \gamma_c t + u_{cpt} \quad \text{(A2)}$$

That is, the model includes pair-period fixed effects, county fixed effects, as well as a county-specific time trend. The coefficient of interest is α, which measures the effect of duration on employment solely using variation from 2013q4 onward (i.e., from no earlier than the quarter immediately prior to expiration). The other independent variable, the log of benefit duration in periods prior to 2013q4, soaks up the effect of duration up to 2013q3; this ensures that, after taking out county fixed effects and county-specific linear trends, the model is comparing employment differences in 2013q4 to employment differences in all quarters in 2014.

The first column of the top panel of **Table A3** shows HMM's estimate of this specification over the 2010q1-2014q4 period, as well as our replication. They estimate a coefficient of -0.0190, with a p-value of zero (to three decimal places) from a block bootstrap procedure. To place this in the context of our other estimates, this would translate into a -1.05 percentage point reduction in EPOP from a 26-to-99 week expansion of duration. While this is smaller than the corresponding estimate in HKMM, it is still substantial, representing about one third of the EPOP drop of the Great Recession; it would also imply

[50] The former correspond to models discussed in Sections 3 through 5 of HMM and the latter correspond to models discussed in Section 6 of HMM.

[51] We understand that HMM takes the spatial difference across pairs manually; as discussed above, this is equivalent to including a full set of pair-period fixed effects.

that the expiration of EUC was responsible for increasing employment in 2014 by about 2 million jobs. When we estimate this equation using the LAUS data that they use on the county pairs in our sample, we estimate a very similar coefficient of -0.0200, with an analytical standard error (clustered at the state-pair level) of 0.0082,[52] which implies a p-value of about 0.015.[53] However, since HMM accessed their data, the entire LAUS series has been redesigned by the BLS, largely to incorporate information from the American Community Survey rather than the Decennial Census.[54] The second column of the first panel shows our estimate from the same specification but with employment derived from the revised data. The coefficient falls in magnitude by three quarters to -0.0048 and becomes statistically indistinguishable from zero. Thus, when using the most recent version of the LAUS employment series, this specification no longer finds that the 2014 EUC expiration caused an employment boom.

HMM also estimate this model using log employment derived from the QCEW and find a modestly negative estimate of -0.0100. In our scale, this would translate to an EPOP effect of -0.558 percentage points from a 26-to-99 week expansion. When we estimate their model we obtain a similar coefficient of -0.0078, corresponding to an EPOP effect of -0.435.[55] While -0.558 is more negative than our 2014 IV specification (-0.024 in the full BCP-FE sample, or -0.094 in the PTT-trimmed sample), the difference is at the bottom end of the range of estimates that can be generated using QCEW data from robustness checks on our main specifications. In results available upon request, we re-estimate our baseline 2014 BCP-FE IV specification using all combinations of the following specification choices: (1) using EPOP, log EPOP, or log employment as the dependent variable,[56] (2) using duration in logs or in levels as the independent variable of interest, (3) keeping county pairs involving North Carolina or dropping them, (4) defining the instrument based on changes in duration immediately upon the EUC expiration, or defining it based on the change between average duration in 2013q4 and the average duration in 2014, (5) starting the sample in 2013q1 or 2013q4, and (6) using seasonally-adjusted or not-seasonally-adjusted data. After translating each estimate to its implied effect on EPOP in levels, we find that these 96 estimates range between -0.637 and 0.473. The EPOP-equivalent estimate from HMM specification using QCEW data (either -0.558 using their estimate or -0.435 using our replication) is within that range, though at the negative end. Furthermore, as with the

[52] In our baseline specifications, we cluster two-way at the state and state-pair level in order to account for any common state-level shocks (including mechanical correlation of errors for those counties that border multiple states). For the sake of this reconciliation exercise, we cluster at the state-pair level. Clustering at the two-way level in this specification increases the standard error to 0.0097.

[53] Our baseline sample includes 1,161 county pairs, and we drop an additional two pairs due to missing data in this specification. While our baseline specification studying the 2014 EUC expiration drops pairs in which either county is in North Carolina, we do not drop such pairs in this reconciliation exercise. HMM report using 1,175 pairs with full data. Such a discrepancy could arise due to reasonable differences in interpretation regarding, e.g., whether counties that touch only on a corner should be included as a "county pair."

[54] See http://www.bls.gov/lau/lauschanges2015.htm for details. We downloaded the current LAUS data on November 10, 2016.

[55] In our baseline specifications in this paper, we seasonally adjust the QCEW data as described in the text. For the sake of this reconciliation exercise, we use not-seasonally-adjusted data.

[56] We do not estimate a specification using employment in levels.

LAUS specification, we find a lower level of statistical precision than HMM: our standard error of 0.0068 would mean that HMM's point estimate of -0.0100 would not be statistically distinguishable from zero at conventional levels.

HMM repeat their analysis with two variants of their benchmark model. First, they replace the county fixed effects and linear trends with interactive effects (Bai 2009) and estimate the model over the 2005q1-2014q4 period. Second, they add to the benchmark model county-specific coefficients on three aggregate time series: the price of oil, aggregate construction employment, and reserve balances with the Fed system. We show these estimates in Panels 2 and 3, respectively, of **Table A3**. The first column shows HMM's estimate and our replication using the pre-redesign LAUS data.[57] These estimates are qualitatively similar to the estimates from the benchmark model. And, like the benchmark model, the coefficient estimates come much closer to zero when post-redesign LAUS employment data is used, consistent with the null effect of benefit expansions that we find in our baseline specifications. We have not been able to replicate their results with the QCEW.

Additionally, HMM estimate "event study" specifications, as described in their Section 6. These specifications are designed to compare employment in 2014 to what is predicted to have occurred in the absence of the EUC expiration based on pre-expiration data. These predictions are formed by estimating a model using data solely from 2005q1 to 2013q4, and by using the resulting parameter estimates to project the future path of employment in a given county. To estimate the pre-event model, HMM regress county-level log employment on county fixed effects, date fixed effects, a county-specific cubic in the quarterly date, and four lags of log employment. They then define their dependent variable e^*_{ct} as the difference between actual log employment and predicted log employment based on the model parameters. Finally, they recover the effect of the EUC expiration by estimating the following model using observations only from 2014:

$$e^*_{ct} = \alpha\big(\ln(D_{c,2014}) - \ln(D_{c,2013q4})\big) + \nu_{pt} + \epsilon_{cpt} \tag{A3}$$

They estimate a coefficient of approximately -0.02, both using employment from LAUS and from the QCEW, meaning that counties which saw larger declines in benefits than their neighbors (i.e., whose independent variable is more negative) experienced higher growth of log employment in 2014, relative to their neighbors, relative to the prediction of their model. As with the estimates found in the "interaction term" models using pre-revision LAUS, this estimate would imply that the 26-to-99 week expansion would explain about one third of the EPOP drop during the Great Recession.

While we have not been able to replicate their results exactly, we do obtain qualitatively similar results.

[57] We calculate standard errors in Panel 2 via a block bootstrap at the state-pair level. We use four factors, as HMM report using for LAUS employment, throughout Panel 2.

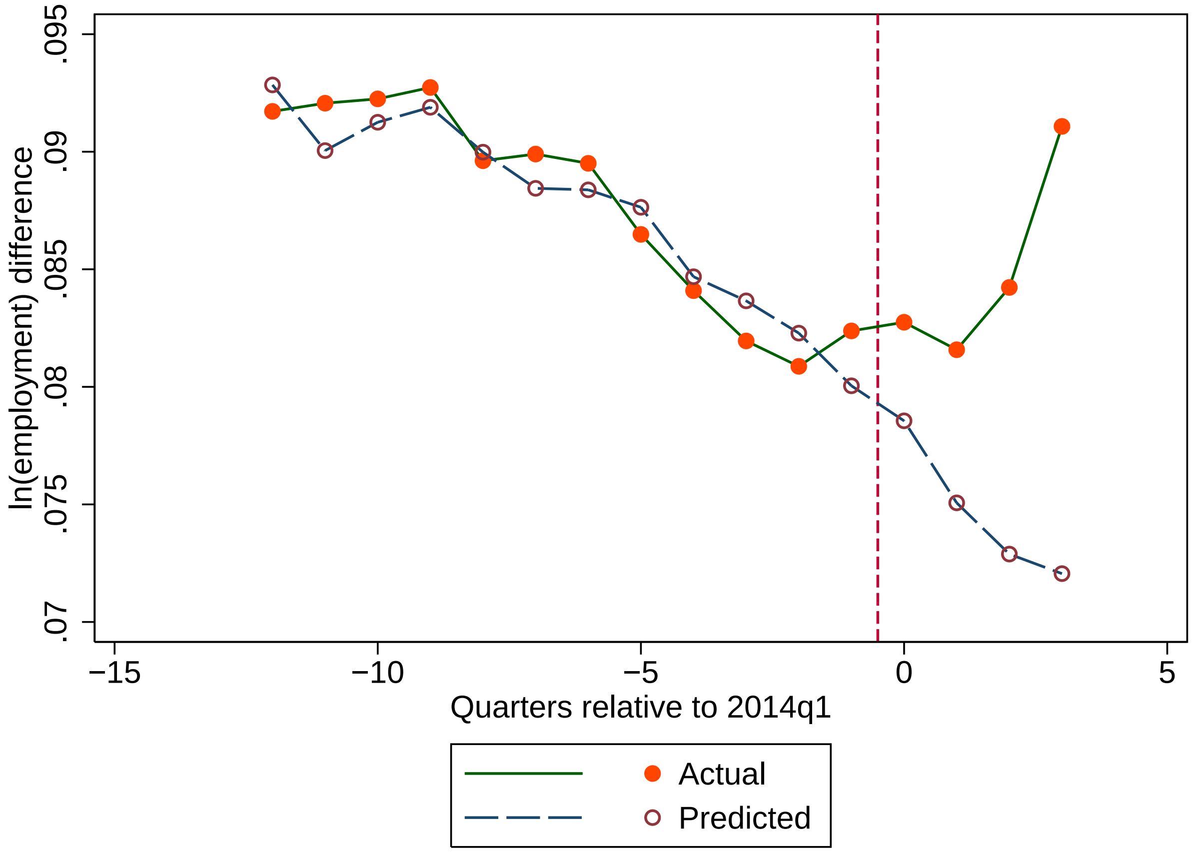
The main result from the event study strategy can be seen immediately in **Appendix Figure A1**, which plots the time series of the average value of log employment, as well as the series of predicted log employment, for high-benefit counties relative to low-benefit counties (where "high" and "low" status is defined by the size of the drop in log duration between 2013q4 and 2014, relative to the county pair partner). The model predicts that employment in high-benefit counties will continue to fall in 2014 relative to their lower-benefit neighbors, when in fact, a modest reversal occurs. The event study approach attributes this to the effect of the EUC expiration. As in the "interaction term" models discussed above, the redesign of the LAUS series affects the results substantially. When we repeat the analysis using the revised data, we find that the coefficient estimate becomes slightly (and insignificantly) positive, as shown in **Appendix Figure A2**. HMM also estimate the event study with QCEW data, and find an estimate of -0.0236, which is larger (in magnitude). When we estimate this model using employment from the QCEW, we find a coefficient of -0.0126 (with a standard error of 0.0113), which is in between our estimates for the specifications with revised and vintage LAUS log employment, respectively.[58] This is shown graphically in **Appendix Figure A3**.

When translated to a change in EPOP, our replication of HMM's event study estimate using the QCEW (-0.703) is substantially more negative than our estimates using EUC expiration, which ranged between -0.024 (full BCP-FE sample) and -0.182 (PTT-trimmed sample). HMM's event study strategy estimates a negative effect of EUC expiration using QCEW data because it constructs a counterfactual where the employment differential between the high and low treatment counties is expected to become more negative in 2014. However, this HMM counterfactual is largely driven by a county-specific polynomial time trend, whose identification is heavily reliant on employment changes that occur up to nine years before the treatment event.[59] As an indication of the type of problem with such a parametric strategy, the employment reversal (both in the QCEW data and, in fact, in the pre-revision LAUS data as well) appears to begin a few quarters prior to the expiration of EUC—a "pre-reversal" which casts doubt on the plausibility of a continuing downward trend as the appropriate counterfactual. In contrast, we take a much more flexible approach by showing whether the employment rates were following parallel trends prior to 2014 by treatment status on the two sides of the border in our 2014 expiration IV. We find that they were, indeed, following parallel trends—as shown clearly in **Figure 9** for the full set of border county pairs. And that this employment gap between the two sides of the border remained largely unchanged following the 2014 expiration. We think the more transparent evidence from the 2014 event that we provide in **Figure 9** raises questions about the causal import of the parametric model used by HMM to construct the counterfactual employment path.

[58]This standard error takes the parameters of the model estimated in the pre-change period as non-random, likely causing us to understate this standard error. HMM use a bootstrapping procedure to construct these standard errors.

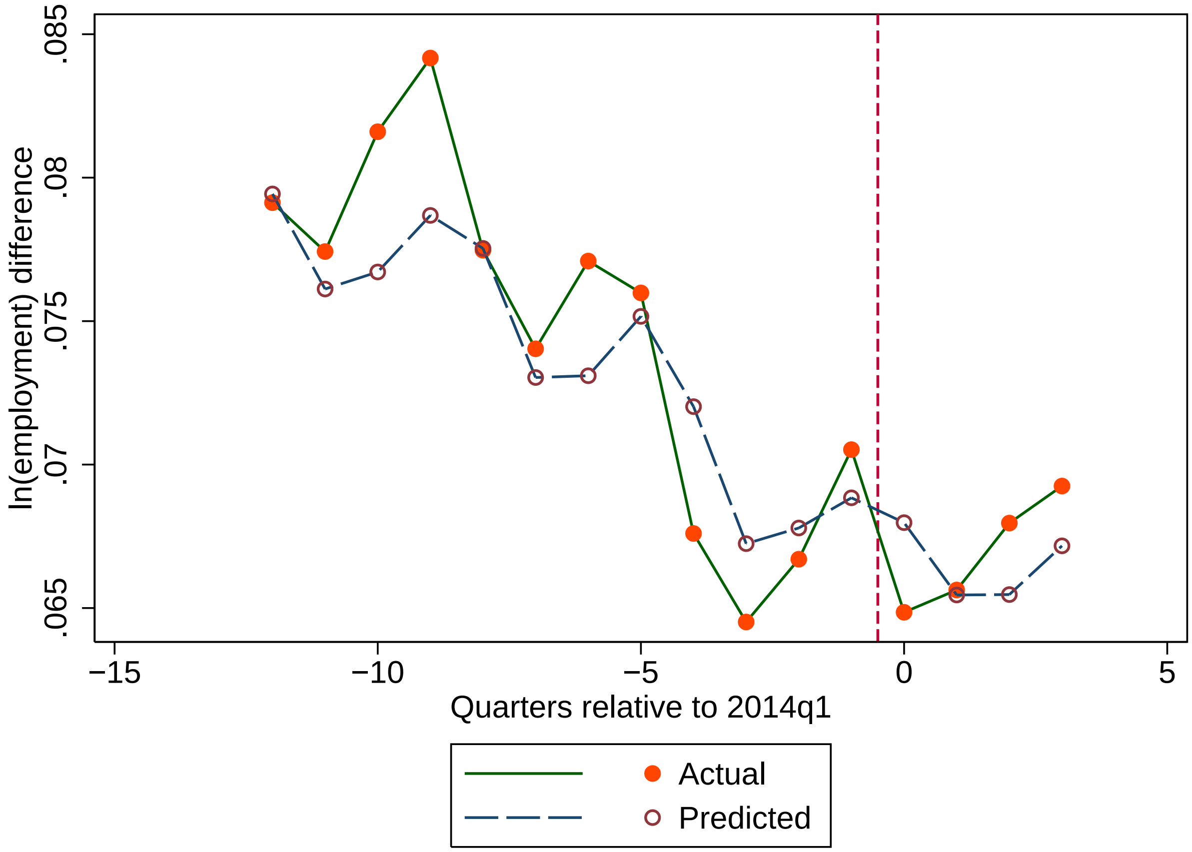
[59]The use of a cubic trend, rather than some other degree of polynomial, does not affect these results substantially.

Figure A1: Replication of HMM event study: Pre-revision LAUS employment



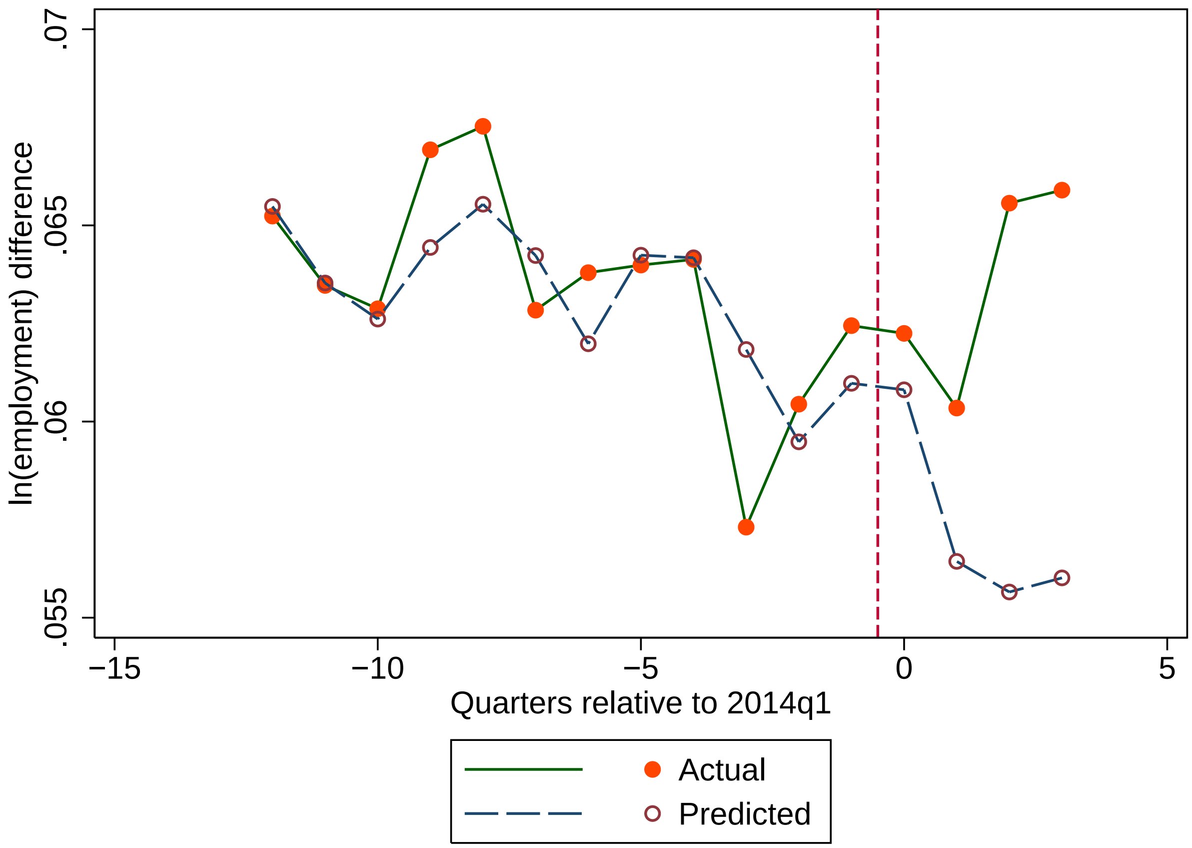
Notes: This figure plots (solid line, solid points) the average difference in log employment between "high" and "low" counties, where a "high" county is defined to have experienced a larger drop in log duration between 2013q4 and 2014 than its neighbor; pairs which experienced identical drops in log duration are not included. The figure also plots (dashed line, hollow points) the average difference in predicted log employment between high and low counties, where the prediction is computed by regressing (on quarterly data from 2005q1 through 2013q4) county log employment on four lags of log employment, time fixed effects, and a county-specific cubic function of the date. Predictions in 2014q1 through 2014q4 are computed recursively. This figure uses employment data from LAUS, prior to the March 2015 redesign.

Figure A2: Replication of HMM event study: Post-revision LAUS employment



Notes: This figure plots (solid line, solid points) the average difference in log employment between "high" and "low" counties, where a "high" county is defined to have experienced a larger drop in log duration between 2013q4 and 2014 than its neighbor; pairs which experienced identical drops in log duration are not included. The figure also plots (dashed line, hollow points) the average difference in predicted log employment between high and low counties, where the prediction is computed by regressing (on quarterly data from 2015q1 through 2013q4) county log employment on four lags of log employment, time fixed effects, and a county-specific cubic function of the date. Predictions in 2014q1 through 2014q4 are computed recursively. This figure uses current LAUS data.

Figure A3: Replication of HMM event study: QCEW employment



Notes: This figure plots (solid line, solid points) the average difference in log employment between "high" and "low" counties, where a "high" county is defined to have experienced a larger drop in log duration between 2013q4 and 2014 than its neighbor; pairs which experienced identical drops in log duration are not included. The figure also plots (dashed line, hollow points) the average difference in predicted log employment between high and low counties, where the prediction is computed by regressing (on quarterly data from 2015q1 through 2013q4) county log(employment) on four lags of log employment, time fixed effects, and a county-specific cubic function of the date. Predictions in 2014q1 through 2014q4 are computed recursively. This figure uses employment data from QCEW.

Table A1: Decomposition of difference between estiamtes from HKMM and BDGK into contributing factors

Step	From HKMM	To BDGK	Average Marginal Effect
Base Estimate	-2.7238***	0.4299	
	(0.6636)	(0.4946)	
No QFD	2.6883***		1.3156***
	(0.6311)		(0.4192)
Align sample	0.8460	1.4613*	0.8629**
	(0.6930)	(0.8803)	(0.3409)
Urate to EPOP	1.3562	1.1334**	0.7421*
	(1.1691)	(0.4869)	(0.4023)
Bai to FE	0.8300	0.6469	0.2186
	(0.7012)	(0.5636)	(0.2968)
Logs to levels		0.0777	0.0146
		(0.3348)	(0.1392)

Notes: The first row reports the total effect of the expansion of UI from 26 to 99 weeks, in percentage points of EPOP, implied by the coefficient estimates of HKMM (column 1) and the full sample BCP-FE estimates of this paper (BDGK) (column 2). The remaining estimates in the first column represent the increased total implied effect of UI when one specification change is made from the original HKMM estimate. The remaining estimates in the second column represent the effect of taking each final step to arrive at the BDGK estimate. Because the total implied effect is not well motivated by theory when using quasi-differenced data in levels, we leave two cells blank in these first two columns. The third column represents the average incremental effect of taking each step along all possible transition paths between HKMM and BDGK estimates, except that we discard transition paths that involve estimating models with quasi-differenced data in levels. See text for details regarding each step and the conversion of each coefficient estimate into an effect on EPOP. Standard errors are calculated via a block bootstrap at the state-pair level with 300 replications. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Transitioning from HKMM to BDGK estimates: Contribution of factors along three particular paths

	Path 1			Path 2			Path 3	
	Coefficient	EPOP effect		Coefficient	EPOP effect		Coefficient	EPOP effect
HKMM reported result	0.0490***	-2.5885***						
HKMM replication	0.0519***	-2.7238***	HKMM replication	0.0519***	-2.7238***	HKMM replication	0.0519***	-2.7238***
	(0.0093)	(0.6564)		(0.0093)	(0.6564)		(0.0093)	(0.6564)
Eliminate QD	0.0086	-0.0355	Urate to EPOP	-0.0025	-1.3676	Align sample	0.0153***	-1.8778***
	(0.0321)	(0.1327)		(0.0020)	(1.0434)		(0.0030)	(0.4642)
Bai to FE	0.1304***	-0.5825***	Elimate QD	-0.0021	-0.1220	Elimate QD	0.0061	-0.0251
	(0.0415)	(0.2021)		(0.0054)	(0.3180)		(0.0224)	(0.0925)
Urate to EPOP	-0.0275**	-1.6064**	Logs to levels	-0.0298	-0.0298	Logs to levels	0.3197*	-0.2046*
	(0.0123)	(0.7034)		(0.2440)	(0.2440)		(0.1692)	(0.1083)
Align sample	0.0059	0.3523	Align sample	-0.2170	-0.2170	Bai to FE	1.0995***	-0.7035***
	(0.0081)	(0.4870)		(0.1405)	(0.1405)		(0.2473)	(0.1582)
Logs to levels (BDGK)	0.4299	0.4299	Bai to FE (BDGK)	0.4299	0.4299	Urate to EPOP (BDGK)	0.4299	0.4299
	(0.4662)	(0.4662)		(0.4662)	(0.4662)		(0.4662)	(0.4662)

Notes: This table presents three transition paths from HKMM's estimates to the full sample BCP-FE estimates of this paper (BDGK). Each cell presents the coefficient estimate, as well as the implied total effect of the 26-99 week expansion of UI expressed as an implied impact of EPOP, in percentage points. Once a step is made in a given path, it is retained in subsequent specifications in the same path. See text for details regarding each step. Standard errors for specifications involving the Bai (2009) interactive effects estimator are calculated via a block bootstrap at the state-pair level with 300 replications. Standard errors for other specifications are clustered twoway at the state and state-pair level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Estimates using the HMM interaction-term model: Alternative data sets and specifications

	(1) LAUS (orig.)	(2) LAUS (rev.)	(3) QCEW
Benchmark			
HMM's estimate	-0.0190***		-0.0100***
	[0.000]		[0.050]
Our estimate	-0.0200**	-0.0048	-0.0078
	(0.0082)	(0.0060)	(0.0069)
Observations	46440	46440	46440
Interactive Effects			
HMM's estimate	-0.0233***		-0.0121***
	[0.000]		[0.030]
Our estimate	-0.0231**	-0.0050	-0.0031
	(0.0099)	(0.0073)	(0.0086)
Observations	92720	92720	92880
Natural Factors			
HMM's estimate	-0.0144***		-0.0141***
	[0.000]		[0.020]
Our estimate	-0.0138	-0.0013	-0.0065
	(0.0104)	(0.0070)	(0.0067)
Observations	46440	46440	46440

Notes: This table reports estimates of α from HMM's "interaction-term" model: $e_{ct} = \kappa[\ln(D_{ct})\mathbb{1}(t \leq 2013q3)] + \alpha[\ln(D_{ct})\mathbb{1}(t \geq 2013q4)] + \nu_{pt} + \epsilon_{cpt}$, under different characterizations of the error term ϵ_{cpt}. In each panel, the top row reports the estimates reported by HMM, with p-values (from a block bootstrap at the state-pair level) in brackets. The second row reports our replication, with standard errors in parentheses. The first column uses log employment from LAUS, prior to the 2015 redesign. The second column uses post-redesign LAUS data, downloaded on September 9, 2016. The third column uses (not-seasonally-adjusted) log employment from the QCEW. The first panel represents the "benchmark" specification, in which $\epsilon_{cpt} = \mu_c + \gamma_c t + u_{cpt}$. The second panel replaces the fixed effects and county-specific trends with interactive effects (Bai (2009)): $\epsilon_{cpt} = \lambda_c' F_t + u_{cpt}$. The third panel adds to the benchmark specification county-specific coefficients on three national time series: the price of oil, employment in the construction industry, and reserve balances with the Fed system. Standard errors in the first and third panel are analytical, clustered at the state-pair level. Standard errors in the second panel are derived from a block bootstrap at the state-pair level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Online Appendix B: Additional Tables and Figures

Figure B1: Increase in UI benefit duration from the November 2008 expansion of EUC

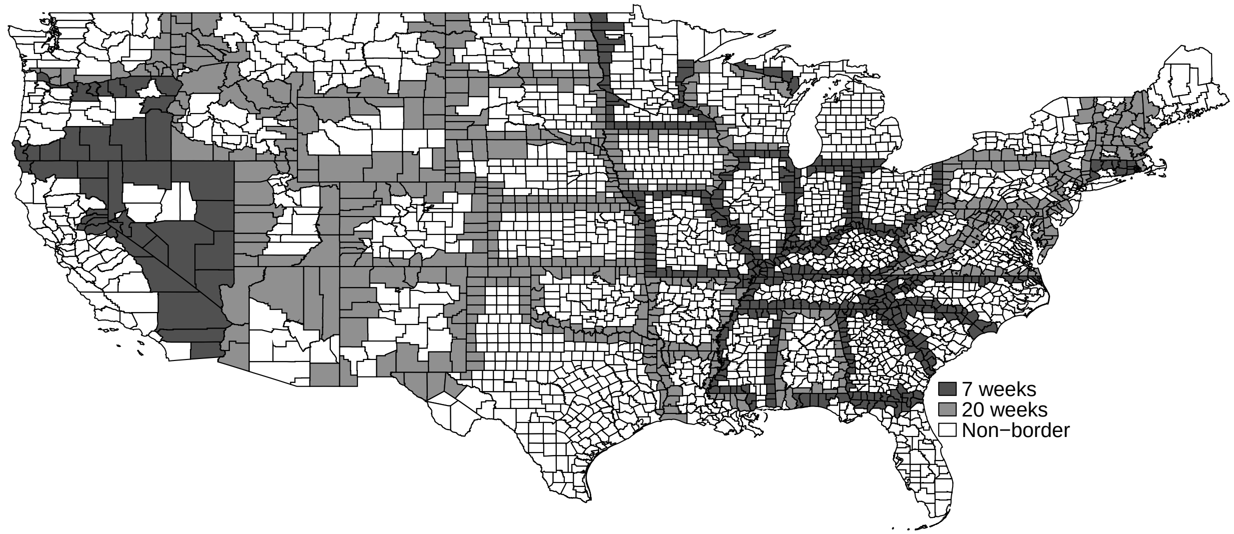
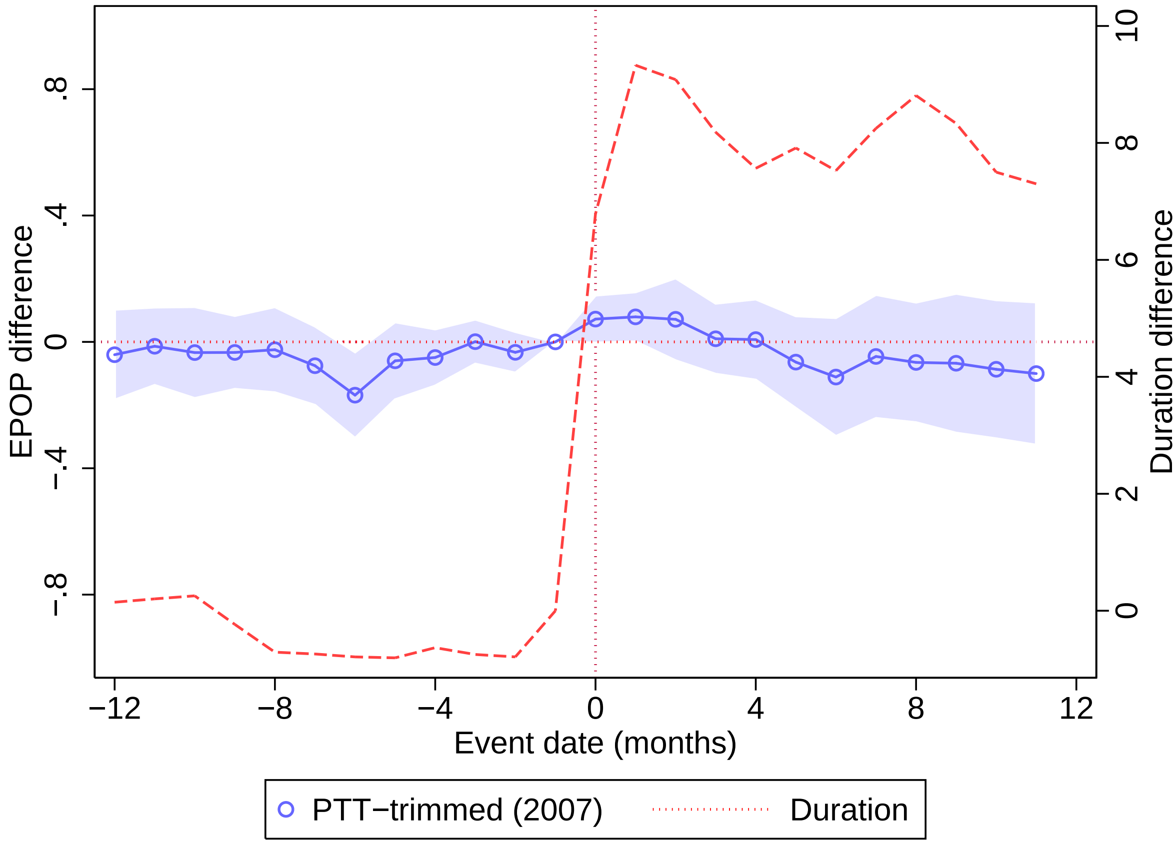


Figure B2: Evolution of average EPOP difference and UI benefit duration difference across state borders: Pooled 2008 expansion and 2014 expiration of EUC, sample trimmed based on PTT estimated through 2007m10



Notes: This figure reports the monthly cumulative response of EPOP (left axis, hollow circles) from the pooled 2008 and 2014 samples, using an alternative trimmed sample. In this figure, PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2007 (not October 2008, as before). See notes to Figure 7 for additional information.

Figure B3: Distribution of EUC differences across border county pairs immediately prior to EUC expiration

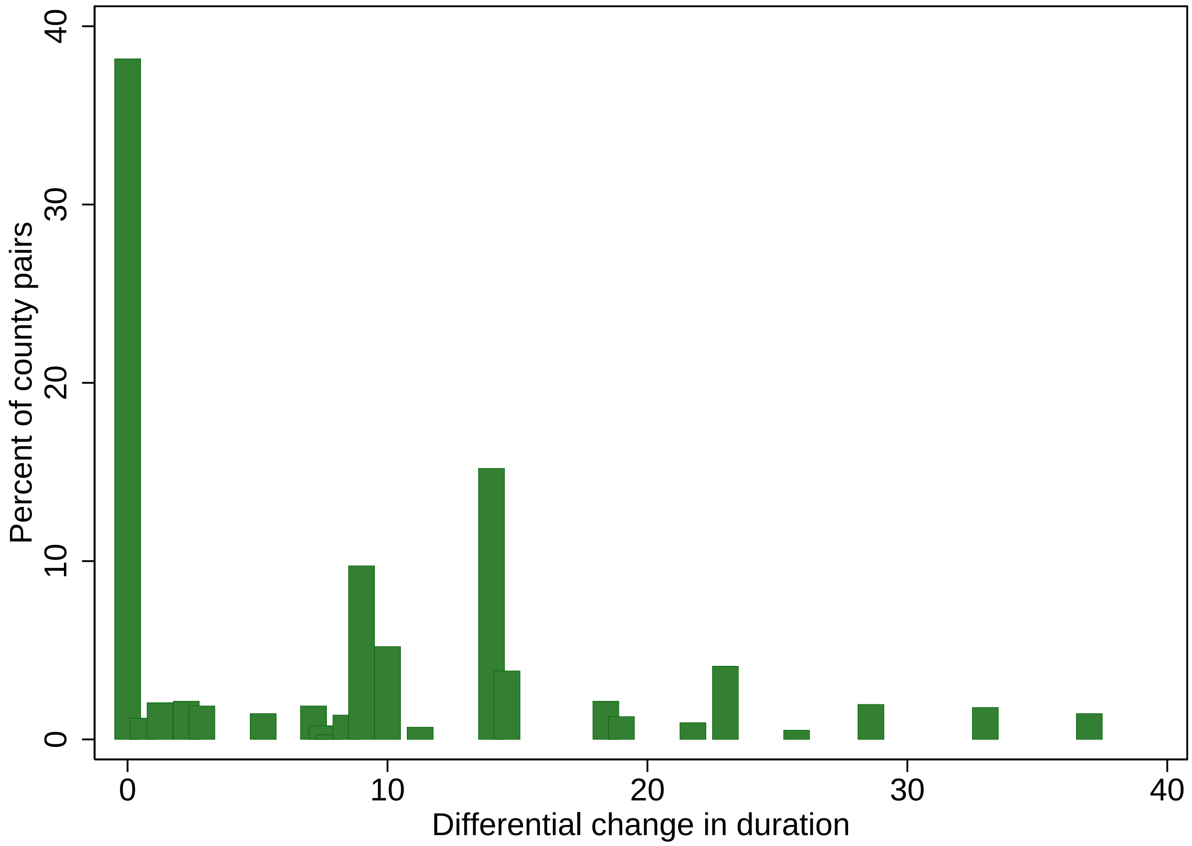
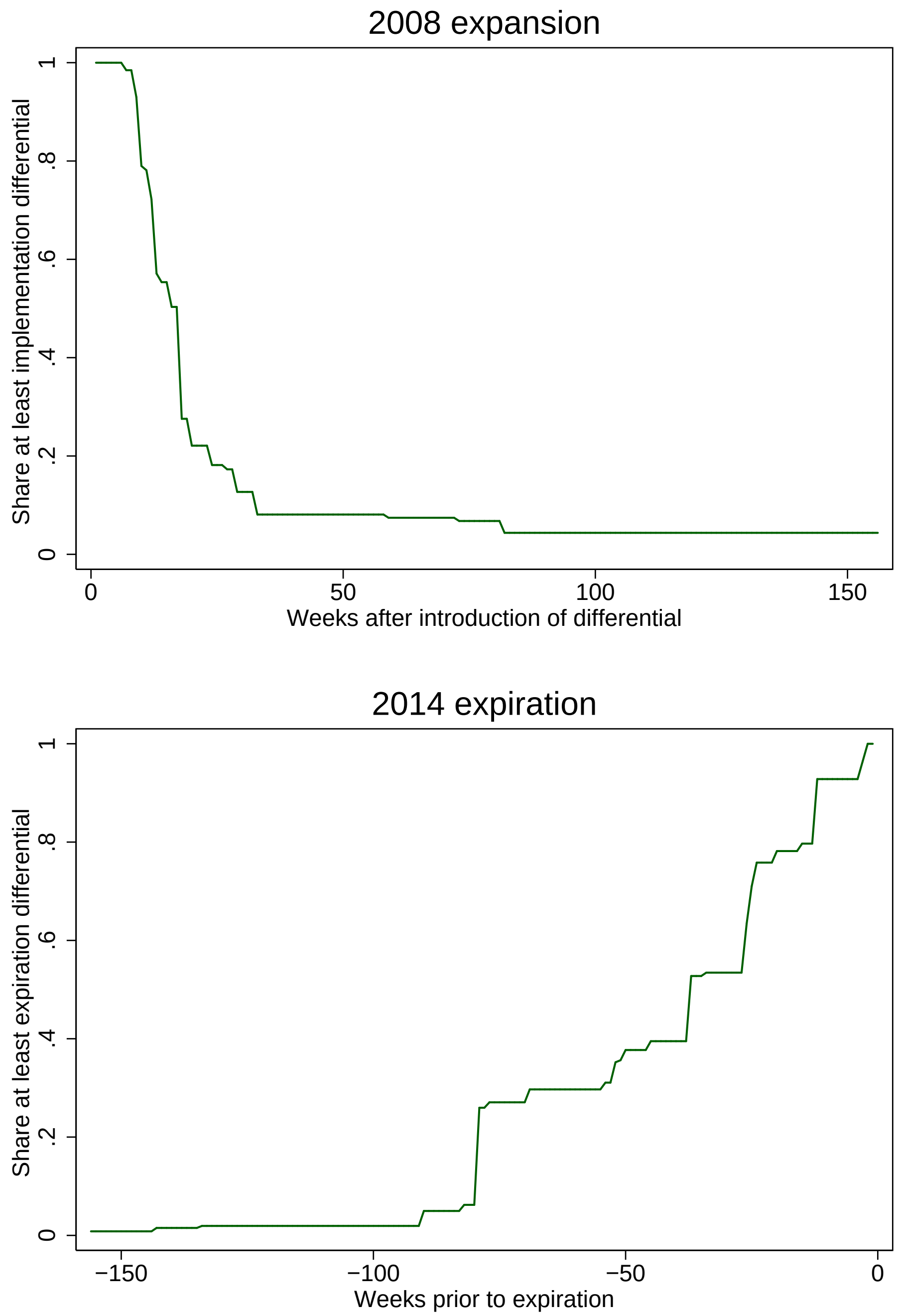


Figure B4: Persistence of duration differences in 2008 and 2014 events



Notes: The top graph plots the share of county pairs that continuously have a duration difference at least as large as immediately after the implementation of UCEA in November 2008. The bottom graph plots the share of county pairs that continuously have a duration difference (moving backward in time) at least as large as immediately prior to the 2014 expiration of EUC. The sample of pairs is restricted to those with differential duration at the time of the event in question.

Table B1: Summary statistics for all counties, all county border pairs, and PTT-trimmed sample of county border pairs

	All counties		**Border counties**		**PTT-trimmed**	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
EPOP (2007)	44.19	18.33	44.51	16.20	42.73	15.17
Private EPOP (2007)	34.58	17.45	34.88	15.47	33.24	14.78
LAUS unemployment rate (2007)	4.857	1.686	4.948	1.777	5.046	1.795
Population age 15+ (2007)	76818.0	243398.5	72692.4	178383.3	55884.2	120677.9
Share white (2005-2009 ACS)	0.796	0.190	0.812	0.181	0.817	0.187
Share black (2005-2009 ACS)	0.0885	0.144	0.0834	0.145	0.0884	0.154
Share hispanic (2005-2009 ACS)	0.0755	0.128	0.0620	0.101	0.0540	0.0961
Share high school grad, less than Bachelor's (2005-2009 ACS)	0.564	0.0665	0.568	0.0640	0.570	0.0610
Share Bachelor's degree or higher (2005-2009 ACS)	0.187	0.0852	0.184	0.0818	0.178	0.0785
Median household income (2005-2009 ACS), 2009 dollars	43299.6	11419.7	42949.1	11725.8	41847.9	11682.7
Newly acquired mortage debt per capita (2007)	3.535	3.216	3.508	3.120	3.216	2.829
Share in cities 50k+ (2010 census)	0.186	0.333	0.188	0.328	0.160	0.304
Minimum weeks of UI eligibility over sample period	23.78	4.365	24.17	4.040	24.20	4.025
Maximum weeks of UI eligibility over sample period	91.37	12.15	90.74	12.38	91.00	12.13

Notes: If a border county appears in j county-pairs in the sample in question, then it appears j times for the purpose of creating estimates in this table. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008.

Table B2: Estimated effect of UI benefit duration on EPOP in specifications without pair-period fixed effects

	(1) All counties	(2) Border counties
No fixed effects	-3.037***	-3.244***
	(0.556)	(0.756)
County fixed effects	-1.826***	-1.768***
	(0.107)	(0.135)
Time fixed effects	-9.550***	-10.670**
	(3.551)	(4.241)
Twoway fixed effects	-0.385	-0.382
	(0.355)	(0.361)
Observations	266944	199692
No. of counties	3104	1129

Notes: This table reports estimates of the form $E_{ct} = \beta D_{ct} + \text{FE} + \epsilon_{ct}$, where E_{ct} is the ratio of employment to population aged 15+, scaled in percentage points. Row 1 considers models without fixed effect. Rows 2-4 consider models with different sets of fixed effects (FE). Standard errors are clustered at the state level in column 1, and two-way at the state and state-pair level in column 2. If a border county appears in j county-pairs, then it appears j times when creating the estimates in column 2.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Additional robustness checks on the effects of UI benefit duration on EPOP: 2008 and 2014 event samples

	2008 sample IV		**2014 sample IV**	
	(1) BCP-FE	(2) PTT-Trimmed	(3) BCP-FE	(4) PTT-Trimmed
1. Baseline	0.549	0.198	-0.024	-0.182
	(2.515)	(1.265)	(0.562)	(0.521)
2. Private EPOP	1.097	0.805	-0.164	-0.214
	(2.514)	(1.209)	(0.633)	(0.517)
3. Correlation-trimmed	0.559	-0.077	-0.392	-0.402
	(2.864)	(1.255)	(0.720)	(0.579)
4. ISLT	0.237	0.286	1.206	0.814
	(1.354)	(0.937)	(0.851)	(0.768)
5. PTT through 2007m10		0.642		-0.029
		(1.365)		(0.479)
6. Quarterly data	0.787	0.538		
	(2.403)	(1.202)		
7. QWI EPOP (quarterly)	0.110	0.863	0.517	-0.020
	(1.679)	(1.519)	(0.579)	(0.572)
8. Distance trimming	1.482	0.889	-0.229	-0.222
	(2.565)	(1.216)	(0.705)	(0.628)
9. Unbalanced panel	0.511	0.198	-0.002	-0.182
	(2.471)	(1.265)	(0.558)	(0.521)
10. Exploit Δ reg. benefits			0.037	-0.139
			(0.555)	(0.508)
11. Drop NC	0.660	0.301		
	(2.808)	(1.406)		
12. Keep NC			-0.437	-0.671
			(0.709)	(0.737)
13. NC: Alt. instrument			-0.037	-0.150
			(0.322)	(0.306)
14. ln($EPOP$)	0.032	0.006	-0.001	-0.008
	(0.044)	(0.023)	(0.009)	(0.010)
	[1.886]	[0.375]	[-0.037]	[-0.436]
15. ln(emp)	0.039	0.011	0.006	0.001
	(0.045)	(0.023)	(0.010)	(0.009)
	[2.327]	[0.622]	[0.320]	[0.068]

Notes: Each cell reports regressions analogous to those reported in Table 3 for the 2008 and 2014 subsamples (each estimated via IV). The estimates in the 1st row correspond to the estimates in panels 2 and 3 of Table 3. The estimates in the 2nd row replace (total) EPOP with the ratio of private employment to population age 15+. In the 3rd row, we trim the set of border county pairs based on the level of correlation between county EPOP and state EPOP over the period 2004m11-2008m10 (see text for details). The 4th row controls for county-specific linear trends. The 5th row trims based on PTT estimated through 2007m10 instead of 2008m10. The 6th row uses quarterly data instead of monthly. The 7th row uses EPOP derived from the QWI (at the quarterly level) instead of the QCEW. The 8th row drops county-pairs whose population centroids are greater than 100km apart. The 9th row includes counties without full EPOP data for each month, which we drop by default. The 10th row uses a modified version of the instrument z_{ct} which exploits all changes in benefits, including changes in regular benefits, which occur at the end of December 2013. Rows 11-13 report estimates using alternative strategies for dealing with North Carolina (NC); by default, border county pairs (BCPs) with one neighbor in NC are kept in the 2008 subsample and dropped in the 2014 subsample. The 11th row completely drops all NC BCPs. The 12th row keeps all North Carolina BCPs. The 13th row keeps NC BCPs but redefines the instrument for NC counties (see text for details). The 14th and 15th row use ln($EPOP$) and ln($employment$), respectively, as dependent variables. The bracketed estimates in these two rows are the level-on-level equivalent, equal to $(\frac{99}{26}^{\hat{\beta}} - 1)\bar{E}$, where $\bar{E}$ is the mean EPOP level in the given sample. Cells which are not applicable in the given sample are left blank. Standard errors are clustered two-way at the state and state-pair level. * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01.

Table B4: Cumulative response of EPOP from distributed lags specification: OLS in first-differences

	(1) BCP-FE		(2) PTT-Trimmed	
	Leads	Lags	Leads	Lags
Contemp.		-0.006		0.034
		(0.096)		(0.084)
Lead/lag 1	0	-0.123	0	-0.063
	(0)	(0.127)	(0)	(0.110)
Lead/lag 2	0.118	0.004	-0.032	0.001
	(0.113)	(0.149)	(0.097)	(0.132)
Lead/lag 3	0.208	0.218	0.083	0.259
	(0.156)	(0.219)	(0.133)	(0.216)
Lead/lag 4	0.263	0.154	-0.047	0.280
	(0.196)	(0.183)	(0.171)	(0.201)
Lead/lag 5	0.148	0.220	-0.075	0.333
	(0.241)	(0.226)	(0.218)	(0.233)
Lead/lag 6	0.243	0.069	-0.037	0.273
	(0.265)	(0.286)	(0.233)	(0.256)
Lead/lag 7	-0.030	0.239	-0.321	0.341
	(0.278)	(0.284)	(0.234)	(0.259)
Lead/lag 8	0.058	0.113	-0.199	0.259
	(0.313)	(0.314)	(0.248)	(0.250)
Lead/lag 9	0.056	0.229	-0.186	0.139
	(0.319)	(0.333)	(0.260)	(0.267)
Lead/lag 10	0.307	0.165	0.122	0.179
	(0.329)	(0.372)	(0.253)	(0.297)
Lead/lag 11	0.403	0.112	0.107	0.106
	(0.341)	(0.372)	(0.247)	(0.312)
Lead/lag 12	0.228	0.168	-0.046	0.136
	(0.334)	(0.390)	(0.249)	(0.336)
Lead/lag 13		0.154		0.198
		(0.399)		(0.346)
Lead/lag 14		0.094		0.056
		(0.406)		(0.355)
Lead/lag 15		0.193		0.230
		(0.465)		(0.409)
Lead/lag 16		0.341		0.360
		(0.440)		(0.386)
Lead/lag 17		0.273		0.304
		(0.463)		(0.402)
Lead/lag 18		-0.003		0.054
		(0.494)		(0.410)
Lead/lag 19		0.155		0.159
		(0.505)		(0.411)
Lead/lag 20		0.162		0.192
		(0.542)		(0.437)
Lead/lag 21		0.168		-0.002
		(0.574)		(0.456)
Lead/lag 22		0.210		0.127
		(0.591)		(0.472)
Lead/lag 23		0.181		-0.030
		(0.577)		(0.476)
Lead/lag 24		0.340		0.012
		(0.590)		(0.502)

Notes: This table reports cumulative monthly lags and leads estimated on the full sample (2007m11-2014m12), using all border county pairs (BCPs) (column 1) and the subset of BCPs in the PTT-trimmed sample (column 2), where all independent variables are divided by 73. The dependent variable is the first-differenced seasonally adjusted ratio of total employment to population age 15+, scaled in percentage points. The regression includes 24 lags and 11 leads and is estimated using EPOP data from 2007m11-2014m12 (and thus duration data from 2005m11-2015m11) in first differences. The zeroth cumulative lag is equal to the estimated coefficient on contemporaneous duration. The jth cumulative lag is equal to the estimated coefficient on contemporaneous duration plus the sum of the estimated coefficient on the 1st through jth lag term. The jth cumulative lead is equal to the sum of the estimated coefficient on the 1st through the $j-1^{th}$ lead term. The 1st cumulative lead is normalized to zero. PTT-trimming removes the quartile of county pairs with the highest differential in linear trends between November 2004 and October 2008. Standard errors are clustered two-way at the state and state-pair level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.