Excess Capacity and Heterogeneity in the Fiscal Multiplier: Evidence from the Recovery Act

Arindrajit Dube[*], Thomas Hegland[†], Ethan Kaplan[‡] and Ben Zipperer[§]

November 9, 2018

Abstract

We estimate local multipliers using cross-county variation in expenditure in the ARRA. We use within-state variation, and include other demographic controls as well as a predicted employment control using an industry shift-share measure. We find that counties receiving more stimulus expenditures had followed parallel employment trends prior to the ARRA as compared to other counties. We estimate an average annualized employment multiplier of 1.211 job-years per $100K spent per county resident. We find strong evidence of heterogeneous treatment effects: the employment response is much greater in counties hit harder by the Great Recession, and hence with likely greater excess capacity. In below median excess capacity counties, the employment multiplier is 0.39. In above median excess capacity counites, the multiplier rises to 2.83. These findings imply that an employment-maximizing stimulus package targeted to high excess capacity counties would have created 83% more (3.60 million) jobs. While our findings are consistent with state-dependent fiscal multipliers, the heterogeneity is not due to the zero-lower bound—since our cross-sectional variation in excess capacity holds the interest rate constant. Instead, our findings suggest that the spatial variation in multipliers reflects variation in the depth of the recession across different labor markets. Consistent with the evidence on hysterisis, we find that the employment impact of the stimulus was long lasting and have likely persisted through the current exapansion.

[*]University of Massachusetts Amherst, NBER and IZA
[†]University of Maryland at College Park
[‡]University of Maryland at College Park
[§]Economic Policy Institute

We thank Joshua Budlender, Owen Thompson and Ruchita Tiwary for excellent research assistance and the Institute for New Economic Thinking for generous financial support. In addition, we thank Emi Nakamura and Jon Steinsson for helpful discussions at an early stage of the project. Any mistakes are our own.

1 Introduction

> We do not have a good measure of the effects of fiscal policy in a recession because the methods that we use to estimate the effects of fiscal policy—both those using the observed outcomes following different policies in aggregate data and those studying counterfactuals in fitted model economies—almost entirely ignore the state of the economy and estimate “the” government multiplier, which is presumably a weighted average of the one we care about—the multiplier in a recession—and one we care less about—the multiplier in an expansion. Notable exceptions to this general claim suggest this difference is potentially large. Our lack of knowledge stems significantly from the focus on linear dynamics: vector autoregressions and linearized (or close-to-linear) dynamic stochastic general equilibrium (DSGE) models. Our lack of knowledge also reflects a lack of data: deep recessions are few and nonlinearities hard to measure”Parker 2011.

There is considerable variation in existing empirical estimates of fiscal multipliers. Some are in the 0-0.5 range (Barro and Redlick 2011, Conley and Dupor 2013); others are near to 1 (Ramey 2011) and yet others are well above 1 (Blanchard and Perotti 2002). Theoretical estimates also vary from near zero (Baxter and King 1993) to well over 1 and sometimes even over 2 (Chodorow-Reich 2017; Christiano, Eichenbaum, and Rebelo 2011; Woodford 2011). Part of the reason for the disagreement, as Parker 2011 points out, may rely on the heterogeneity in the multiplier as a function of the degree of slack in the economy. However, what is of greatest interest for macroeconomic stabilization purposes is the multiplier during recession. Unfortunately, estimation of the multiplier as a function of excess capacity has been elusive. The reasons for this are three-fold. First, at a country level, the identification must exclusively rely on time series variation. However, the timing of expenditures is correlated with the business cycle itself and thus expenditures are confounded by the state of the economy. Second, there is a limited sample size from which to perform statistical inference.

Some recent efforts have attempted to use a narrative approach in order to estimate fiscal multipliers (C. D. Romer and D. H. Romer 2010). However, this approach is plagued by small sample sizes, which is particularly problematic when investigating heterogeneous effects of fiscal stimulus by the amount of slack in the economy. Third, when excess capacity is high, interest rates tend to be low and fiscal expenditure more effective. Thus, the interest rate confounds the estimate of the heterogeneity in the multiplier. This is particularly important when excess capacity is quite large because the interest rate on government debt is likely to be close to the zero lower bound on nominal interest rates and a large recent literature argues that multipliers are much higher when policy interest rates are sticky downwards (Christiano, Eichenbaum, and Rebelo 2011; Eggertsson et al. 2003; Eggertsson 2011; Woodford 2011).

Beginning with Auerbach and Gorodnichenko 2012, a burgeoning literature has attempted to estimate the heterogeneity in the multiplier using time series variation (Auerbach and Gorodnichenko 2012, Baum, Poplawski-Ribeiro, and Weber 2012, Clemens and Miran 2012, Fazzari, Morley, and Panovska 2015, Ramey 2011; Ramey and Zubairy 2018; Mittnik and W. Semmler 2012; W. Semmler and A. Semmler 2013). However, it is focused on the difference in the multiplier at the zero lower bound in nominal interest rates. One recent paper (Ramey and Zubairy 2018) also uses time series variation and employs vector autoregression methods to estimate the differential multiplier in recessions versus booms. They, however, try to separate out the impact of being at the zero lower bound in interest rates versus the impact of having high excess capacity. They find no differential due to high degrees of excess capacity but no difference in in multipliers across periods of boom or bust. Their multiplier estimates lie between 0.4 and 0.8 throughout during recessions and booms and during periods of high interest rates as well as low interest rates. However, the time series based estimation is not well identified. Moreover, estimates are not stable to timing mis-specification as well as method of construction of impulse response functions (Ramey and Zubairy 2018).

An alternative approach to multiplier estimation using national time series is estimation using intra-national variation in fiscal expenditures over time to estimate multipliers (Chodorow-Reich et al. 2012; Chodorow-Reich 2017; Conley and Dupor 2013; Feyrer and Sacerdote 2011; Moretti 2010; Nakamura and Steinsson 2014; Serrato and Wingender 2016; Shoag et al. 2010). There is even a small set of papers estimating local multipliers using variation in spending across areas in the American Recovery and Reinvestment Act or 'Obama Stimulus Bill' (Chodorow-Reich et al. 2012; Conley and Dupor 2013; Feyrer and Sacerdote 2011).

Our paper uses cross-county variation in expenditure during the ARRA to estimate fiscal multipliers. We semi-parametrically estimate the multiplier as a non-parametric function of the degree of excess capacity. We show that our estimates of the fiscal multiplier satisfy time placebos, are robust to inclusion of Bartik controls for evolution of employment based upon the sectoral composition of local employment and national trends in employment by sector. We also instrument for the amount spent with an instrument for the 'shovel-readiness' of federal contracts in the county and find similar results. Moreover, Boone, Dube, and Kaplan 2014 show that allocation of funds in ARRA was not correlated with the unemployment rate[1]. Our paper is closest in topic to Ramey and Zubairy 2018 in that we estimate the heterogeneity in the multiplier as a function of excess capacity, partialing out interest rate effects. It is closest to Feyrer and Sacerdote 2011 in terms of methods. However, they do not estimate differential multipliers by excess capacity. In contrast to Ramey and Zubairy 2018, we find a five-fold increase in the fiscal multiplier for above-median excess capacity counties compared to below-media excess capacity counties. One recent paper (Michaillat 2014) shows that when unemployment is higher, public sector employment crowds out fewer matches and thus the multiplier is countercyclical. Our methods are unable to distinguish whether the multiplier is higher in high excess capacity areas because in those areas there is more idle capital (Keynes 2018), whether there is less crowdout of matching efficiency from employment programs (Michaillat 2014), or whether in areas of high excess capacity, unemployment is also greater and consumption multipliers are therefore higher (Gross, Notowidigdo, and Wang 2016).

Of course, as with the rest of the local multiplier literature, our use of spatial variation comes at a cost. We do not directly estimate a national fiscal multiplier. Moreover, because expenditures are financed through federal rather than through local taxes, what we are estimating is more akin to a transfer multiplier in an open economy as opposed to a national fiscal multiplier. On the positive side, in addition to the improved identification, estimating locally allows us to say something about the channels through which the multiplier works (Farhi and Werning 2016). One standard mechanism through which the multiplier is thought to work is that reduced wealth from the increased debt obligations needed to finance the fiscal expansion increases labor supply (Baxter and King 1993). However, this channel should be shut off in our local estimates as long as those receiving more and less funds should not pay on average higher taxes from the increased expenditure.

[1] These results are at the Congressional District level but they hold at the county level as well.

Since we show that our results are robust to controlling for deciles of average tax obligations per capita, we would expect that to the degree that public and private expenditures are substitutes, the counties with higher public expenditure should have lower employment. In fact, we find the opposite to be true.

In addition to providing an estimate of the multiplier as a function of excess capacity, our paper also adds to the literature in a few other ways. First, we try to better bridge the gap between the local multiplier and national multiplier literature by showing the effect of spatial aggregation. We do not find that the multiplier changes much as we spatially aggregate. This is in contrast to Nakamura and Steinsson 2014 who find smaller multipliers when going from the state level to the national level. We interpret the difference as due to lowering the degree of spatial measurement error and estimating over a less open (more aggregated) economy. However, we find little evidence that the greater multipliers are driven by a greater degree of endogeneity of expenditure at a higher level of spatial aggregation.

Second, we provide more evidence on the dynamics of expenditure, computing non-parametric impulse responses similar to what is estimated in the Vector Autoregression literature (Blanchard and Perotti 2002, Ramey 2011, Ramey and Zubairy 2018). Allowing more non-parametric estimation of impulse responses to fiscal expenditure is beneficial because it allows for arbitrary non-linearities in the time path of the effect of expenditure. This is a particularly important contribution given Ramey and Zubairy 2018's demonstration of the sensitivity of the multiplier to the method of constructing impulse response functions in non-linear models such as the VAR models pioneered by Auerbach and Gorodnichenko 2012 to estimate heterogeneity in the multiplier over the business cycle.

Finally, we point out that, given the heterogeneity which we find in the fiscal multiplier, our estimates reflect an average multiplier which is averaged over each dollar spent. We separate out the economic and political aspects of the multiplier by computing the multiplier for a politically unconstrained government which optimally targets federal dollars to the highest multiplier areas. In other words, in addition to estimating the actual multiplier, we compute what the multiplier would have been had it been optimally spent based upon the information that the government would have had access to at the time.

In section 2 of this paper, we present our empirical methodology. In section 3, we describe the data that we use in estimating our results. In section 4, we present our results and finally, in section 5, we conclude.

2 Empirical methodology

To estimate the effects of fiscal stimulus we regress county-level quarterly employment and earnings on a quarterly measure of county-level stimulus and additional controls. We primarily use fixed effect methods. To explore heterogeneity of the multiplier in the extent of stimulus, we add a quadratic term in the amount of stimulus. To explore heterogeneity of the effects of stimulus across counties of varying excess capacity, we estimate over split samples as well as estimate using the semiparametric smooth coefficient estimator proposed by Li et al. 2002.

2.1 Primary specifications

Our primary regressions are county-level panel and county-level fixed effects regressions[2]. Letting i,s, and t denote, respectively, geographic region[3], state or division, and quarterly indices, the panel specification

$$Y_{it} = \alpha + \beta S_{it} + \gamma B_{it} + F_{st} + tD_{it}^{'}\Delta + \epsilon_{it} \tag{1}$$

regresses quarterly employment or earnings per capita outcomes Y_{it} on stimulus per capita S_{it}. Controls in this specification include a Bartik shift-share control for predicted employment based upon industrial shares of employment in county i in 2008 Quarter 1, state-time-specific fixed effects F_{st} and demographic controls which vary over time and across counties denoted byD_{it}. Demographic controls are Census 2000 estimates of percents black, Hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. These are all multiplied by linear time trends. In some specifications, the fixed effects are at the stateXtime level or at just the pure time level: F_t.

2.2 Time Aggregation

We also present dynamic estimates including lagged effects of stimulus. We present cumulated effects

[2]We also ran long difference regressions of the change in the employment to population ratio on stimulus from 2008Q3-2011Q3. These results look similar to the fixed effects results which use the full panel. Therefore, we report only the panel results. Long difference regression results are available from the authors upon request.

[3]For most specifications, the geographical region is county. However, in some specifications, regions are a higher level of spatial aggregation than county.

over time from our dynamic regressions. In particular, we estimate:

$$Y_{it} = \alpha + \sum_{k=0}^{8} \beta_k S_{it-k} + \gamma B_{it} + F_{st} + tD_{it}^{'}\Delta + \epsilon_{it} \tag{2}$$

We then report contemporaenous effects: β_0, impulse response incorporating contemporaneous effects and two quarter lagged effects $\sum_{k=0}^{2} \beta_k$, and other impulse responses cumulated up to two years of lagged effects: $\sum_{k=0}^{8} \beta_k$. Stimulus awards begin to flow in mid-2009. Our panel specification begins in 2008q1 and ends in 2011q3.

2.3 Heterogeneous effects

We also consider specifications where the effects of stimulus depend on the extent of stimulus. Denoting all right-hand side variables except for stimulus as the vector $\mathbf{Z}_{it}$, then we estimate

$$Y_{it} = \alpha + \beta_1 S_{it} + \beta_2 S_{it}^2 + \gamma B_{it} + F_{st} + tD_{it}^{'}\Delta + \epsilon_{it} \tag{3}$$

This differs from our main specification only by the addition of a quadratic term on the amount of stimulus spent in a quarter within a county.

We additionally examine how the extent of county-level excess capacity alters the effects of stimulus. As we discuss in section 3, our measure of excess capacity E_i is based on pre-period industry shares and is constant for each county. First we break down counties into above median excess capacity and below median excess capacity and estimate (1) separately for above-median and below-median excess capacity counties. We then consider a more flexible non-linear interaction or semi-parametric smooth coefficient model:

$$Y_{it} = g(E_i) + (S_{it}, \mathbf{Z}_{it})'\mathbf{h}(E_i) + \epsilon_{it} \tag{4}$$

where the scalar g and vector $\mathbf{h}$ are unspecified functions of excess capacity. We estimate equation 4 at each excess capacity percentile e^p by linear regressions of Y_{it} on S_{it} and $\mathbf{Z}_{it}$ for observations whose population-weighted kernel-based distance is near e^p, as suggested by Li et al. 2002. We county-cluster bootstrap these estimates to conduct inference.

2.4 Spatial Aggregation

In our final empirical specification, we aggregate spatially to estimate multipliers in economic geographic units with a lower degree of openness and a higher percentage of expenditures within the geographic area. For each county, we add employment, stimulus expenditure and demographics for all counties whose population centroid is within a given radius of an observation's base county. We consider 30, 60, 90 and 120 mile radii specifications. We then estimate equation (1) with each county's data replaced by the aggregated data. Of course, this induces mechanical serial correlation. We correct for this autocorrelation by using our knowledge of the spatial dependency under the maintained assumption that underlying county data is not inherently spatially correlated. In particular, we use Conley and Dupor 2003 standard errors in a panel context, which allows arbitrary serial temporal and spatial autocorrelation of the error term for all counties whose centroids are within 2*D miles of each other, when we consider aggregation by D miles. Thus, the selector matrix is not block diagonal as it would be in the clustering case.

3 Data

The primary data sources are the Quarterly Census of Employment and Wages (QCEW) from the Bureau of Labor Statistics (BLS) [4] and the recipient-reported American Recovery and Reinvestment Act award data from recovery.gov (Stimulus).[5] Additionally we utilize a variety of other demographic and geographic data as control variables from multiple data sources which we detail below and we employ tools to implement spatially-based regressions.

3.1 Outcomes

The key dependent variables are employment and earnings (wage bill) per capita. We construct our main employment measure, EPOP, by dividing county-level employment reported at the quarterly level from the QCEW and dividing it be intercensal population estimates from the Bureau of the Census for the population

[4] https://www.bls.gov/cew/

[5] This data was formerly available at \url{http://www.recovery.gov/FAQ/Pages/DownloadCenter.aspx}; a portion of it is currently available at https://www.nber.org/data/ARRA/

aged 15-64.[6]. Quarterly earnings data are taken directly from the county-level BLS QCEW. We did the aggregate wage bill in a county-quarter by the size of the 15-64 aged population. We also then divide by 100,000 so that wage bill impacts can be interpreted as impacts upon per capita wages per $100,000.These data are not seasonally adjusted. To calculate quarterly population levels we use the annual July 1 intercensal estimates published by the US Census Bureau[7] as our third quarter population estimates and interpolate estimates among quarters assuming a quarterly geometric growth rate.

3.2 Treatment

The treatment variable which we use in the paper is the amount of stimulus funds per capita spent in a county in a quarter. We construct this variable using recipient-reported stimulus award data which we downloaded from www.recovery.gov. "These data are a panel of individual contracts, grants, and loans reported quarterly beginning in 2009q4 through 2013q3, though we only use data on contracts and grants. Award data is also reported for a single 2009q1-2009q2 period, which we we assign to 2009q2.

Recipient-reported data is available for prime awardee recipients and sub-recipients who receive more than $25,000. Prime awardees report the overall award amount and sub-recipients report subawards. We construct our dataset using prime awards and their award amount. We then add in subrecipients and their subawards.

Prime awards report their expenditure-to-date on a quarterly basis. We use this data to construct prime awards' expenditure per quarter. Subawards do not report expenditure-to-date, instead reporting only when the subaward is active. We assume subawards are spent at the same rate as their prime awards during the time period when the subaward is active. Note that a reasonably small number of prime awards report nonmonotonicities in their cumulative amount spent over time, such as when they report cumulative expenditure levels of 0 in their final reporting periods. In cases like these where awards report cumulative

[6]The definition of EPOP used by BLS divides aggregate employment in the CPS by the 16+ population. Our definition differs from this definition in a few respects. First, we use the QCEW rather than the CPS. The QCEW is based on unemployment insurance records reported to state governments by firms and then transmitted to the U.S. Census. It differs from the CPS in that it is a census, not a sample. Thus, we have accurate measures of employment by county in each quarter. However, though it contains 98% of jobs, it does not contain the self-employed. Additionally, at the county level, we use July intercensal estimates of population for 15-64 (http://www.census.gov/popest/) rather than 16+. Our measures of employment are thus smaller than those used by BLS to construct national EPOP. However, our measures of population are larger. Overall, our EPOP measure is smaller than the national measure constructed by BLS.

[7]http://www.census.gov/popest/

expenditure levels lower than the prior cumulative expenditure level for up to two consecutive quarters, we correct these nonmonotonicities by replacing those observations with data linearly interpolated based on the two quarters surrounding the block of nonmonotonic cumulative expenditure levels. When expenditure levels fall in an award's final reporting period, we use the value from the prior period. In the rare cases that these corrective procedures fail to produce monotonically increasing cumulative expenditure data, we drop the data from our sample. "

In the raw recovery data project status can sometimes behave nonmonotonically. Awards may regress in status, or awards fail to be reported at all during some periods. Reported status-unadjusted award amounts may furthermore change from quarter-to-quarter, presumably due either to reporting error or to total award amounts legitimately changing over time. For example, it is not necessarily the case that the status-unadjusted total award amount for a prime recipient reported in 2009q3 will be the same amount reported in 2011q3. One reasonable restriction is to create a dataset R excluding awards with award status regressions and reporting breaks, and using only the most recent (2011q3) reported unadjusted award amount and recipient geography. We also create a dataset $\tilde{R}$ that does not impose these time restrictions: it does discriminate on status or reporting breaks, and it uses the contemporaneous status-unadjusted award amount and geography reported during that quarter.

For a given award in the recipient-reported data, monies awarded and place of performance zip code are available for prime recipients and subrecipients. We assign zip codes to counties using the MABLE/Geocorr2K Census 2000 zip code-to-county crosswalk. [8]In the rare cases where place-of-performance zip code data is not reported or is reported with error, we use an award's reported city and state to assign it to a county. When this data is also not available, we use the award recipient's zip code.

We divide our stimulus measure by the 15-64 aged population from the Census. We then also divide by $100,000. Thus our estimates of the impact of stimulus can be interpreted as the impact of additional $100,000 of expenditure per capita in a county[9].

[8]When a zip code in multiple counties, we allocate awards based on population shares using MABLE/Geocorr2K Census 2000 population allocation factors.

[9]Since we divided the dependent and independent variables by the same measure of population, any measurement error in the county-level intercensal population estimates produced by the Census introduces negative bias into our estimated multiplier.

3.3 Excess Capacity

We also create an excess capacity variable so that we can estimate the differential impact of stimulus in high versus low excess capacity areas. We compute a county's excess capacity as the absolute value of the largest observed one year reduction in industry shift-share predicted employment in the county from 2006 or 2007 to 2008. Intuitively, this measure ranks counties by the size of the reduction in the employment to population ratio that their industry-composition suggests they should have received, prior to passage of ARRA.

3.4 Controls

Because our identification using panel data relies on county fixed effects, we do not use lagged outcomes as controls for fear of biasing our OLS estimates. Instead we use as a control predicted employment and earnings, using pre-period county-level industry shares and contemporaneous national level employment to predict actual employment in the manner of Bartik (1991). Specifically, we first calculate county-level average overs the years 2006 and 2007 employment (earnings) shares of national employment (earnings) at the three-digit NAICS level. Then we multiply these county-NAICS shares by contemporaneous national three-digit NAICS employment. We sum the resulting county-NAICS series over NAICS categories to form a single, time-varying predicted employment (earnings) series for each county.

In addition to Bartik-predicted outcomes and geographic and time dummies, we also employ a variety of pre-period demographic controls in the hope of increasing the precision of estimates and also to account for some of the selection bias remaining in stimulus assignment. We use US 2000 Census county-level estimates of percents black, Hispanic, urban, and under poverty, as well as county-level median income.[10] Because of the central role of housing wealth in the most recent recession, we also use two county-level housing variables derived from the loan origination reported under the Home Mortgage Disclosure Act (HMDA): the 2006 average value of home purchase loans, and the 2006 total of all HMDA loans divided by county population.[11]

Our regressions are panel data regressions county fixed effects. Thus, we simply interact these demographic controls with a time trend. For spatially-based regressions we calculate neighbors' and neighbor distance bins

[10] http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/13402/ascii
[11] http://www.ffiec.gov/hmda/hmdaproducts.htm

by defining the distance between two counties as the distance between their population-weighted centroids.[12] The population-weighted centroids are taken from the Missouri Census Data Center MABLE/Geocorr2K database[13], which uses Census 2000 geography and population data.

4 Results

We begin by showing the amount spent over time. Figure 1 panel A shows the amount spent nationally over time. The status-adjusted amounts increase slowly over time. The majority of the funds were spent by 2012. However, even at the end of our data set, 13% of the had not been allocated. First, however, we show that money was spent continuously throughout time. The percent of status-adjustment stimulus money spent is roughly linear in time up to 2013 as is the average dollar amount per capita spent. Moreover, we see in Figure 2 that the amount spent was relatively randomly distributed across the United States. Looking at the shaded map of amount spent by county, we see no obvious spatial patterns of expenditure. To effectively use the variation in expenditure over both space and time, we use a panel approach at the county and quarterXyear level.

4.1 Own-county Multipliers

We present our baseline estimates of the contemporaneous own-county impact of stimulus in Table 1. These estimates are broken down into two super-columns. The left super-column contains estimates of the impact of stimulus on own-county employment and the right super-column contains estimates of stimulus on the own-county wage bill. In the first row, we show static estimates of current stimulus expenditures on employment in the following quarter. In the next 4 rows, we show the cumulative sums of lagged coefficients for models with 2, 4, 6, and 8 lags respectively. All these models have a symmetric number of leads. Finally, in the bottom row, we show cumulative leads for the 8 lag model. We show four separate specifications in each super column with progressively more stringent sets of controls. All columns contain county fixed effects. The first column additionally contains quarterXyear (henceforth time) fixed effects. The second column puts

[12] We implement the distance calculation in Stata using Kenneth L. Simons' \texttt{circnum} command: http://homepages.rpi.edu/~simonk/technical.html

[13] http://mcdc2.missouri.edu/websas/geocorr2k.html

in a set of controls which vary across counties and over time: a Bartik predicted outcome variable [14] percent black, hispanic, urban, and under poverty, median income and 2006 average home purchase price for loans and 2006 total HMDA loans per capita. The third column drops demographic and economic covariates but replaces time fixed effects with stateXtime fixed effects. The final column re-adds the demographic and economic controls to the model with stateXtime and county fixed effects. The effects are largely stable across specifications

We use our estimates with the most stringent set of controls as our main estimates. We feel that they are the best identified and also they generally have the greatest precision. Our results do vary some across specification. However, they are all of the same sign, they are all statistically significant at a 95% level or higher, and they are all within fifty percent of our preferred estimate. This is true for both the employment estimates as well as the wage bill estimates. Our benchmark estimate of the effect of an additional $100,000 of stimulus expenditure per capita upon EPOP is that the additional expenditure increases EPOP by 1.211 percentage points. Since ARRA was $787 billion and the U.S. population in 2011 (when the median ARRA dollar was spent) was 311 million, this amounts to $2500 per person or 2.5% of $100,000 per person. Thus, our estimates imply that ARRA expenditures raised EPOP by 3.06 percentage points. Restricting to the contracts, grants and loans component to ARRA, expenditures were $308 billion. They were slightly more than 1/3 of the total size of ARRA. Restricting just to the contract, grants and loans which we use in our analysis of the impact of ARRA, we find that ARRA added 1.20 percentage points of EPOP.

We also estimate the impact of stimulus on wage bill. The 8-lag time-aggregated wage bill coefficient is 0.416. This implies that an extra $100,000 of expenditure per capita raises the per capita wage bill by $41,600. If we divide this by the employment multiplier, we find that if wages for employed people did not rise, the average person employed by the stimulus would have made $34.300. In other words, the marginal jobs created by ARRA were likely lower paying jobs. Overall, our time-aggregated results are consistent with a positive and significant multiplier.

We note that our own-county estimates are larger than the estimates in Feyrer and Sacerdote (2011). This is mostly because we adjust for the timing of expenditure and we include both prime awards and subawardees

[14] When the outcome variable is the employment to population ratio, then the Bartik prediction is for the employment to population ratio; when the outcome variable is the wage, then we compute a Bartik predicted wage. The Bartik controls are computed using county-level employment and wage bill from the QCEW, averaged over the 2006-2007 time period. Predictions are made using industrial composition at the three-digit NAICS level.

when we can place the subawardee expenditures. Without these adjustments, our estimates are quite similar to those by Feyrer and Sacerdote (2011). We thus largely attribute Feyrer and Sacerdote's small estimates to measurement error in the timing of stimulus and to spatial measurement error in the sub-prime awards. Feyrer and Sacerdote (2011) find significantly larger estimates at the state level, consistent with a significant impact of spatial measurement error on estimation.

The sign of our estimates is informative and useful in better understanding the channels through which stimulus worked. It is important to note that we are estimating the impact of expenditures holding fixed tax payments as long as tax payments are not correlated with stimulus expenditures. We empirically estimate but do not report[15] estimates controlling for deciles of tax payments per capita. Thus, we show that our estimates do not reflect differences in future potential tax burden across counties but rather differences in expenditures. Different from traditional national estimates, our local estimates thus effectively hold expected future tax payments constant. Therefore, net wealth increases are at least weakly larger in the areas which receive stimulus. Since, in the baseline macro model, an increase in wealth should reduce rather than increase labor supply, the fact that we estimate positive coefficients indicates that the multiplier is not likely through the wealth-labor supply channel. The leading alternative channel is the Keynesian demand-side channel.

4.2 Placebo Tests

Of course our multipliers may not reflect causal effects of stimulus but rather the differential evolution of EPOP and wages across counties which would have occurred in the absence of stimulus spending. For example, if counties with worse crises received more money but also thus had larger recoveries, it is possible that ARRA expenditures could merely reflect the depth of the crisis as well as the subsequent recovery. In Table 2, we invert our main specifications and regress aggregate stimulus in a county over our sample period in separate regressions on changes in pre-ARRA changes in county macroeconomic measures. In particular, we regress on change in employment between 2006Q1 and 2007Q1 as well as between 2006Q1 and 2009Q1. We do the same with respect to change in the wage bill over the same two time periods. We also regress a measure of the severity in the pre-Great-Recession drop in employment on aggregate stimulus: the maximum quarter-to-quarter pre-Great-Recession dip (over the period 2006Q1–2007Q4 to 2008Q1-2009Q1).

[15] Estimates are available from the authors upon request.

We additionally do the same with predicted Bartik employment. All these regressions control for state fixed effects. We estimate on the full set of counties but estimate on our placebo treatment measure and our measure interacted with a dummy for above median excess capacity.

Out of 12 regressions, we find only one with statistically significant coefficients: stimulus expenditure is significantly predicted by Bartik employment in low excess capacity counties. An additional percentage point decline in predicted EPOP decline is correlated with an additional $50 per capita of ARRA expenditure in the county. When we add our baseline controls to our placebo regressions, the amount falls to $22 per capita though remains statistically significant. The correlation between actual EPOP between 2006Q1 and 2009Q1 (as opposed to Bartik predicted EPOP) is actually somewhat higher. However, the standard errors for Bartik-predicted employment are small due to high serial correlation in the Bartik measure. Recent work has noted that high degrees of serial correlation lead to very low standard errors when Bartik employment is a dependent variable (Borusyak, Hull, and Jaravel (2018)). Nonetheless, for this reason, we control for Bartik employment in our main specifications. Besides the statistically significant coefficients on Bartik employment in low excess capacity counties, the only other statistically significant coefficient is of wage bill per capita in high excess capacity counties and this is only statistically significant at a 10% level and only with baseline controls. We thus find that stimulus expenditures in a county was not well predicted by evolution of employment and wages before the passage of ARRA thus is likely not endogenous to the evolution of employment and wages in a county.

4.3 Heterogeneity by Excess Capacity

In Table 2 and Figure 2, we present estimates of heterogeneity by the degree of excess capacity. Recent macroeconomic theory (Eggertsson (2011); Christiano, Eichenbaum, and Rebelo (2011); Woodford (2011)) suggests strong heterogeneity in the multiplier when the interest rate reaches the zero lower bound. However, others have suggested that government expenditures may vary also by the degree of slack or excess capacity in the economy (Keynes (2018); Parker (2011)). There are many reasons as to why the mutliplier may vary with the degree of excess capacity. In counties with greater slack, employment does not necessarily rely upon large capital investments. In addition, in areas with high unemployment, consumers have may be more liquidity constrained and have a higher propensity to consume out of money spent.

We estimate the impact of stimulus funds heterogeneously across counties with higher and lower excess capacity. We do this in two ways: parametrically by estimating separatley for above median excess capacity counties and semi-parametrically across all counties simultaneously. We begin with our parametric estimates in Table 2. Again, we separate out the presentation of our estimates into two super columns. In the left super column, we present employment multipliers and on the right wage bill multipliers. Each super column contains three columns. These columns show pooled estimates (identical to the estimates in Table 1), estimates from below-median excess capacity counties, and estimates from above-median excess capacity counties.

We find very strong differences between low and high excess capacity in their estimated multipliers. Our cumulative lag estimates for both wage bill and employment in low excess areas are substantially smaller and statistically insignificant. The stimulative effect of an extra $100,000 per capita of government expenditure in below-median excess capacity counties is to raise EPOP by 0.502 per capita over two years. This translates to 1.6 percentage points over two year and a quarter years or 0.8 per year. Following (Chodorow-Reich (2017)), we compute the implicit output multiplier as GDP per worker divided by cost per job. In low excess capacity counties, our employment multiplier estimate translates to an output multiplier of 0.559. This is consistent with other measures of output multipliers in "good times" (Ramey and Zubairy (2018)).

Turning to above median-excess capacity counties, we find a substantially larger employment multiplier. The multiplier is over 5 times the size. An additional $100,000 of expenditure per capita in high excess capacity counties yields 2.546 extra jobs per person. ARRA overall is estimated to have increased employment in high excess capacity areas by 8.24 percentage points over two and a quarter years (or 3.67 per year). This translates into an output multiplier of 2.837. It is important to point out that these cross-sectional multipliers are identified off of differential changes across counties with high versus low excess capacity. Since interest rates were the same in high and low excess capacity areas (and in fact, policy rates were at the zero lower bound for most of the period), our estimates are not confounded by differential multipliers at the ZLB as is common with time series estimates (Ramey and Zubairy (2018)). In addition to the greater ability to estimate with greater identification uncofunded by the state of the economy, the ability to unconfound state-contingent multipliers from interest-rate contingent multipliers is one of the greatest benefits of our cross-county panel estimation strategy.

The wage bill multipliers show similar heterogeneity in patterns. The low excess capacity wage bill coefficient implies an additional expenditure of \$100K per person yields a statiscally insignificant additional \$9K per person in wage income. In high excess capacity counties, an additional expenditure of \$100K per person yields an addition \$107K wage bill per person. This implies that ARRA generated overall an additional \$3.7K per person in high excess capacity counties.

Figure 3 plots the individual lagged effects of stimulus expenditure upon EPOP as well as wage bill in high and low excess capacity counties respectively. In the top panel, we show effects in low excess capacity counties. In these counties, cumulative lags weakly increase over time for both figures. However, cumulative response stops rising after 6 quarters, effects remain low for all quarters, and 95% confidence intervals (shown in red) contain zero in all quarters.

By contrast, in the high excess capacity counties, cumulative effects of stimulus expenditure upon EPOP are statistically distinguishable from zero with a 95% level of confidnece in all quarters. Moreover, cumulative effects rise monotonically over quarters following treatment. Moreover, the last cumulative lag is the highest, suggesting that long term cumulative effects are even higher than in the first two years after stimulus expenditure. This suggests a sizable long-run effect of stimulus expenditures. We also see similar patterns in wage bill effects with the minor exception that the contemporaneous effect is, by itself, not statistically distinguishable from zero with a 95% level of confidence.

Our main effects are estimated using lag operators. However, in a balanced panel, longer lags are estimated off of a truncated sample of time periods. Thus, dynamic estimates can reflect causal effects of treatment or compositional differences in lag estimation. To address this concern, we also estimate the dynamic impact of stimulus by regress EPOP and the wage bill respectively on a set of time dummies interacted with the (time-invariant) average stimulus amount spent in a county. We plot these estimates in Figure 4. The estimates are similar to and thus confirm our main estimates. In low excess capacity counties, we see small and statistically insignificant rises in both EPOP and in the wage bill. In high excess capacity counties, by contrast, we see immediate, large and persistent increase in employment for counties that received greater treatment.

In Figure 5, we also show semi-parametric plots of stimulus upon both EPOP and the wage bill. As we can see from the first panel, the multiplier increases close to weakly montonically in excess capacity. Greater excess capacity yields weakly higher multipliers. The multiplier is mostly flat up until 1 percentage point of excess capacity and then rises. It increases roughly five fold from a two percentage point excess capacity to a 4.5 percentage point excess capacity after which it seems to stabilize. We cannot know for sure as the paucity of observations in that range increases standard errors sufficiently that inference is made difficult. We also see similar patterns in the wage bill multipliers. We find substantial heterogeneity in the multiplier as a function of excess capacity.

4.4 Non-Linear Impacts of Stimulus Funds

One possible explanation for the heterogeneity in the multiplier between high excess capacity regions and low excess capacity regions is that more money was spent in high excess capacity regions and the multiplier increases with the amount spent. In Table 4, we show that this does not explain our heterogeneity results. In particular, we regress the employment to population ratio as well as the wage bill on a quadratic function of amount spent per capita per $100,000. We present results in Table 4. Again, we break our results into pooled effects, effects in low excess capacity counties and effects in high excess capacity counties.

Different from our main results, we estimate our non-linear effects using contemporaneous expenditure only without lags or leads. We do this because the non-linear quadratic model is quite taxing. Our baseline contemporaneous estimates yield an immediate impact of 0.281 jobs per $100,000 spent on average, broken into 0.805 jobs in high excess capacity counties and a statistically insignificant 0.098 jobs in low excess capacity counties. Both the overall and the high excess capacity estimates are statistically significant at a 1% level. We also find similar results for our wage bill estimates though the pooled coefficient is only significant at a 10% level. These estimates are 1/3-1/4 of our 8-lag estimates, consistent with Figure 4 which shows a quick rise after initial expenditures.

We then estimate a quadratic model. The linear component, unsurprisingly, is larger in the quadratic model than in the linear models and the quadratic terms are negative. In the low excess capcity counties, the terms are statistically insignificant. In the pooled counties sample, only the linear coefficient is statistically significant. The linear component is almost twice as large as in the linear model. In high excess capacity areas, the linear coefficient is almost three times as large as in the pooled set of counties. It is also almost twice as large as in the high excess capacity linear model. We find that the marginal impact of stimulus turns negative at approximately $900 per person in a quarter. Stimulus is estimated, in high excess capacity areas, to on average have a negative effect as of approximately $1900 per person in a quarter.

Overall, we do find evidence of declining marginal effect of stimulus within a quarter. However, declining marginal effect of stimulus is inconsistent with high excess capacity counties have higher estimated multipliers because they on average get more money and the multiplier increases in amount spent.

4.5 Sector Specific Multipliers

As evidence that our local multipliers are picking up demand-side effects, we show how fiscal expenditure impacts on employment and wages vary by sectors of the local economy. Funds allocated to public sector schools, to Medicaid and to other public programs are likely to increase public employment. In addition, contracts given to manufacturing firms are likely to increase local employment in manufacturing. However, stimulus expenditures may also increase employment in non-tradable sectors which did not receive federal funds through a consumption multiplier. In this section, we estimate employment multipliers by sector. In particular, we run our main regressions of employment (relative to population) by industry on aggreagte stimulus expenditure. The impact of aggregate stimulus expenditure in the county on industry-specific employment captures both direct contractual effects of stimulus expenditure as well as demand spillovers. We use the QCEW's employment series broken down by industry. Our sample size drops slightly because in a small number of county-quarters, data on employment by industry data is missing for disclosure reasons.

We first break down employment into public sector employment and private sector employment. We estimate effects on all counties, on low excess capacity counties and on high excess capacity counties separately. Sensibly, all coefficients on industry employment are smaller than the coefficient on overall employment. The only sizable and statistically significant coefficient for the pooled sample of counties is that on public sector employment. expands significantly overall from ARRA expenditures. We find that public employment expands by 0.8 jobs per $100,000 spent. However the average effect cloaks dramatic heterogeneity in impact across high and low excess counties.

In low as well as high excess capacity counties, the public sector employment impact is near identical at 0.8 jobs per $100,000 of expenditure. We also see a substantial and statistically significant rise in the wage bill. For every $100,000 per capita, the public sector wage bill rises by approximately $25,000. This is, again, true in both low excess and high excess capacity counties. The remarkable similarity in the coefficients across high and low excess capacity counties are evidence of (1.) the exogeneity of stimulus expenditures conditional upon controls and (2.) the absence of demand spillovers to public sector employment and wages.

We see a coefficient on private sector employment in low excess capacity areas which is equal in size but opposite in sign to the coefficient on public sector employment. It is moreover statistically significant at a 10% level. This negative impact is spread across all private sector industries except for construction though in all cases, coefficients are far from statistically differentiable from zero. We also see statistially significant positive effects on wage bill and negative wage bill effects in private industry in low excess capacity counties. This expansion of public sector employment and contraction of private sector employment suggests that in the absence of slack, workers are drawn from employment in the private sector. We see distinctly different patterns in the high excess capacity counties.

In high excess capacity counties, we see no negative employment or wage bill effects. The point estimates for tradable employment and tradable wage bill are negative but tiny and statistically indistinguishable from zero. In particular, the coefficient for goods production is slightly positive and very small. Since almost none of the federal funds spent locally on tradables will impact local demand for those same tradables, we expect that impacts of expenditure to be relegated to the direct effect of tradable contracts, grants and loans. We thus expect tradable sector multipliers to be low. The public sector jobs multiplier is stable at 0.8 as in the low excess capacity counties. The private sector job multiplier is 1.524; for non-tradables, it is 1.251 and for services, it is 1.159. For every $100,000 spent, there are over 1.5 jobs created in the private sector and 1.2 of those jobs is in the service sector. These coefficients are sizable and statistically significant. In addition, the wage bill coefficients are similarly sizable and statistically significant.

Our industry-specific results are very useful in validating the model and explicating the channels through which expeniture impacts employment and the wage bill. We see substitution from private to public sector employment in the low excess capacity areas and a demand multiplier effect in high excess capacity areas.

4.6 Effectiveness of an optimally allocated stimulus

In this section we compute the number of job-years created from the contracts, grants and loans portion of the Obama stimulus bill. The contracts, grants and loans totaled approximately 1/3 of total stimulus expenditures. We also compute the number of job-years that would have been created had the money been spent solely in above median excess capacity counties.

Our baseline estimates suggest a time-aggregated employment multiplier of 1.211 jobs per $100,000 spent. Since $261 billion were spent on the contracts, grants and loans portion of the ARRA bill, that leaves us with 3.161 million job years created. As shown in Section 4.2 and in Boone, Dube, and Kaplan 2014, the amount of stimulus provided to an area is very weakly correlated with the unemployment rate in the area. Under the assumption that the multiplier is the same in all regions, this allocation would maximize the efficacy of the stimulus. However, in this paper, we have shown that the multiplier in high excess capacity regions is many times greater than the multiplier in low excess capacity areas. This poor targeting of high unemployment areas lowers the average multiplier that we estimate.

We now compute the number of jobs created if the stimulus had been optimally spatially targeted towards high excess capacity counties. We do this using a simple back of the envelope calculation. We assume that there are two multipliers: one for above median excess capacity counties and a separate one for below median excess capacity counties. We also assume that the multipliers do not change with increased expenditures. We do note that even if stimulus had been allocated to only above median excess capacities, excess capacity in the above median counties would still have remained above the median.

We compute the output multiplier following (Chodorow-Reich (2017)) by dividing income per worker by cost per job. We use the year 2011 as a benchmark year since the median ARRA dollar was spent in 2011. In 2011, income per worker was $111,400. Our estimated effect on EPOP per $100K of stimulus expenditures per capita was to create 1.211 jobs per capita. This translates into a cost per job of $82,600. Our employment multiplier thus translates to an output multiplier of 1.349. However, if money had solely been allocated to above-median excess capacity counties, under the assumption that the average multiplier wouldn't have declined with the additional expenditure, the multiplier would have been 2.546 and the cost per job would have been $39,300. The associated output multiplier would have thus been 2.837. The multiplier would have been 110% higher. It would have more than doubled.

This computation is revealing for two reasons. First, from a policy perspective, it shows the importance of optimally allocating funds. Of course, optimal allocation of funds might make bill passage more difficult. However, the consequences pn welfare of allocating funds in a spatially optimal manner are large. Second, this computation makes a methodological point about multiplier estimation. Most macroeconomic theories are divorced from the political economy of bill passage. To the degree that the economy has a single multplier, this is a useful abstraction. In the presence of spatial heterogeneity in the multiplier, estimating national multipliers using national data estimates an average multiplier rather than an employment maximizing multiplier.

The disadvantages of local multiplier estimation are inability to incorporate impacts of monetary policy, difficulties in incorporating effects of increased national debt accumulation, and differential openness of a local area such as a county relative to a country. However, national level multiplier estimation suffers from greater endogeneity problems. In addition, however, as we show in this paper, the internal allocation of funds impacts the national multiplier and thus the national multiplier cannot be used to test purely economic theories of the multiplier.

5 Conclusion

In this paper, we estimate local employment multipliers. We also convert our employment multipliers to fiscal multipliers. On average we find a jobs multiplier of 1.2 which means that an extra $100,000 of expenditure translates into 1.2 extra jobs over 2 years. This jobs multiplier yields an equivalent output multiplier of 1.3. However, this average multiplier masks substantial heterogeneity in the multiplier by excess capacity. We both parametrically and semi-parametrically estimate the multipliers as a function of excess capacity. We find large differentials between low and high excess capacity regions. Even during the Great Recession, we find no statistically significant, cumulative impact of public expenditures on overall employment in counties below median excess capacity. The evidence from these counties is consistent with the additional public employment largely crowding out private employment, leading to an employment multiplier is an additional 0.5 jobs for every $100,000 of expenditures. However, for the counties above median excess capacity, we find a substantially larger contemporaneous employment multiplier of 2.5 jobs per person for every $100,000. This translates into a fiscal multiplier in above median excess capacity areas of slightly more than 2.8.

Our local estimates are larger in the private sector than in the public sector. They are also concentrated in non-tradable industries, particularly in services and construction. Since non-tradables are disproportionately impacted by local demand and tradables are not, this is precisely what we would expect from demand-side stimulus. We also find that multipliers do decline in amount spent. However, these declines are small.

The employment effects we find are perhaps surprisingly pesistent, even through the end of our sample period in 2017. This persistence suggest that the gains we document here may understate the ultimate, long run impact of the policy. These findings are also consistent with the finding of cross-sectional hysterisis documented in Yagan (2018) who find the employment losses during the Great Recession were highly persistent. By reducing the severity of employment loss from the Great Recession, the ARRA likely reduced the extent of scarring in the labor market in hard-hit counties.

Our estimates are useful for understanding when and where public funds are effective at increasing employment and output and for designing employment maximizing stimulus programs. However, the spatial heterogeneity in the multiplier underlies an important point in testing theories of the multiplier. The aggregate national multiplier is influenced by the political economy of the spatial allocation of funds. Thus estimates using national data implicitly test a joint economic and political hypothesis. In contrast, by using spatial variation in the multiplier, it is possible to compute a spatially-optimal employment-maximizing multiplier and test an economic hypothesis.

We hope that future work will improve upon what we have done by better reconciling local multiplier estimates with national estimates. This reconciliation could be improved in three ways. First, we have focused solely upon labor market impacts of stimulus. However, stimulus could impact capital income as well. Second, we have primarily focused on non-tradable sector employment because that is what is easily identifiable using cross-county variation. However, the magnitude of the multiplier would presumably be greater if we could incorporate effects on the tradable sector. There could also be qualitative differences between tradable and non-tradable sector multipliers. If tradable sectors are more likely to use increasing returns to scale technologies, then the multiplier in tradable sectors might display non-linearity in amount spent in contrast to what we have found in predominantly non-tradable sectors. Finally, we have estimated our effects during the Great Recession when the nominal interest rate on government debt was at zero. This presumably makes monetary policy relatively ineffective but fiscal policy quite effective. Our estimates would be improved if they could be generalized outside of the ZLB interest rate zone and also if they could incorporate endogenous responses of monetary policy. We hope future empirical work will make progress in these three ways.

Finally, we also would like to see better theoretical explanations for why the fiscal multiplier is increasing in excess capacity. We can imagine three classes of theories focusing on impacts in the labor market, consumption and production. Michaillat 2014 has a model where the multiplier is decreasing in the degree of labor market tightness. Alternatively, in high excess capacity areas, a higher fraction of individuals are likely liquidity constrained and thus have higher average marginal propensities to consume. More money spent in these areas thus might generate a larger consumption multiplier and thus a larger fiscal multiplier. Finally, it is possible that in higher excess capacity areas, it is easier to hire labor without accompanying capital investments due to slack in the usage of capital. Our empirical results cannot differentiate between a labor market tightness effect from a liquidity effect or an excess capacity effect. Our paper thus calls on future theoretical work on the potential impact of excess capacity on the multiplier as well as empirical work differentiating between different theories of the countercyclical multiplier for reasons unrelated to the zero lower bound in nominal interest rates.

References

Auerbach, Alan J and Yuriy Gorodnichenko (2012). "Measuring the output responses to fiscal policy". In: *American Economic Journal: Economic Policy* 4.2, pp. 1–27.

Barro, Robert J and Charles J Redlick (2011). "Macroeconomic effects from government purchases and taxes". In: *The Quarterly Journal of Economics* 126.1, pp. 51–102.

Baum, Ms Anja, Mr Marcos Poplawski-Ribeiro, and Anke Weber (2012). *Fiscal Multipliers and the State of the Economy*. 12-286. International Monetary Fund.

Baxter, Marianne and Robert G King (1993). "Fiscal policy in general equilibrium". In: *The American Economic Review*, pp. 315–334.

Blanchard, Olivier and Roberto Perotti (2002). "An empirical characterization of the dynamic effects of changes in government spending and taxes on output". In: *the Quarterly Journal of economics* 117.4, pp. 1329–1368.

Boone, Christopher, Arindrajit Dube, and Ethan Kaplan (2014). "The political economy of discretionary spending: Evidence from the American Recovery and Reinvestment Act". In: *Brookings Papers on Economic Activity*, pp. 375–428.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2018). “Quasi-experimental Shift-share Research Designs”. In: *arXiv preprint arXiv:1806.01221*.

Chodorow-Reich, Gabriel (2017). *Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?* Tech. rep. National Bureau of Economic Research.

Chodorow-Reich, Gabriel et al. (2012). “Does state fiscal relief during recessions increase employment? Evidence from the American Recovery and Reinvestment Act”. In: *American Economic Journal: Economic Policy* 4.3, pp. 118–45.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo (2011). “When is the government spending multiplier large?” In: *Journal of Political Economy* 119.1, pp. 78–121.

Clemens, Jeffrey and Stephen Miran (2012). “Fiscal policy multipliers on subnational government spending”. In: *American Economic Journal: Economic Policy* 4.2, pp. 46–68.

Conley, Timothy G and Bill Dupor (2003). “A spatial analysis of sectoral complementarity”. In: *Journal of political Economy* 111.2, pp. 311–352.

— (2013). “The American Recovery and Reinvestment Act: Solely a Government Jobs Program?” In: *Journal of monetary Economics* 60.5, pp. 535–549.

Eggertsson, Gauti B (2011). “What fiscal policy is effective at zero interest rates?” In: *NBER Macroeconomics Annual* 25.1, pp. 59–112.

Eggertsson, Gauti B et al. (2003). “Zero bound on interest rates and optimal monetary policy”. In: *Brookings papers on economic activity* 2003.1, pp. 139–233.

Farhi, Emmanuel and Iván Werning (2016). “Fiscal Multipliersâ: Liquidity Traps and Currency Unions”. In: *Handbook of Macroeconomics*. Vol. 2. Elsevier, pp. 2417–2492.

Fazzari, Steven M, James Morley, and Irina Panovska (2015). “State-dependent effects of fiscal policy”. In: *Studies in Nonlinear Dynamics & Econometrics* 19.3, pp. 285–315.

Feyrer, James and Bruce Sacerdote (2011). *Did the stimulus stimulate? Real time estimates of the effects of the American Recovery and Reinvestment Act*. Tech. rep. National Bureau of Economic Research.

Gross, Tal, Matthew J Notowidigdo, and Jialan Wang (2016). *The marginal propensity to consume over the business cycle*. Tech. rep. National Bureau of Economic Research.

Keynes, John Maynard (2018). *The general theory of employment, interest, and money*. Springer.

Li, Qi et al. (2002). "Semiparametric smooth coefficient models". In: *Journal of Business & Economic Statistics* 20.3, pp. 412–422.

Michaillat, Pascal (2014). "A theory of countercyclical government multiplier". In: *American Economic Journal: Macroeconomics* 6.1, pp. 190–217.

Mittnik, Stefan and Willi Semmler (2012). "Regime dependence of the fiscal multiplier". In: *Journal of Economic Behavior & Organization* 83.3, pp. 502–522.

Moretti, Enrico (2010). "Local multipliers". In: *American Economic Review* 100.2, pp. 373–77.

Nakamura, Emi and Jon Steinsson (2014). "Fiscal stimulus in a monetary union: Evidence from US regions". In: *American Economic Review* 104.3, pp. 753–92.

Parker, Jonathan A (2011). "On measuring the effects of fiscal policy in recessions". In: *Journal of Economic Literature* 49.3, pp. 703–18.

Ramey, Valerie A (2011). "Can government purchases stimulate the economy?" In: *Journal of Economic Literature* 49.3, pp. 673–85.

Ramey, Valerie A and Sarah Zubairy (2018). "Government spending multipliers in good times and in bad: evidence from US historical data". In: *Journal of Political Economy* 126.2, pp. 850–901.

Romer, Christina D and David H Romer (2010). "The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks". In: *American Economic Review* 100.3, pp. 763–801.

Semmler, Willi and Andre Semmler (2013). "The Macroeconomics of Fiscal Consolidation in the European Union". In:

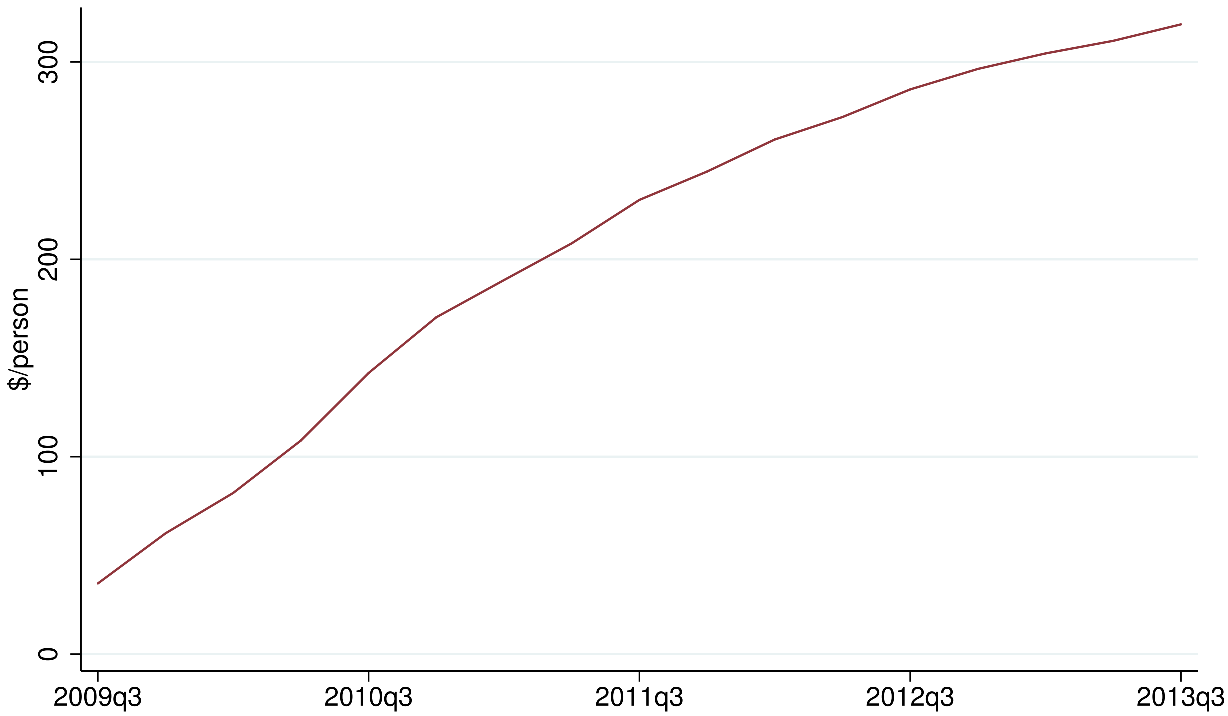
Serrato, Juan Carlos Suárez and Philippe Wingender (2016). *Estimating local fiscal multipliers*. Tech. rep. National Bureau of Economic Research.

Shoag, Daniel et al. (2010). "The impact of government spending shocks: Evidence on the multiplier from state pension plan returns". In: *unpublished paper, Harvard University*.

Woodford, Michael (2011). "Simple analytics of the government expenditure multiplier". In: *American Economic Journal: Macroeconomics* 3.1, pp. 1–35.

Figure 1: Cumulative flow of stimulus awards

Total status-adjusted awards per capita



Proportion of awards spent

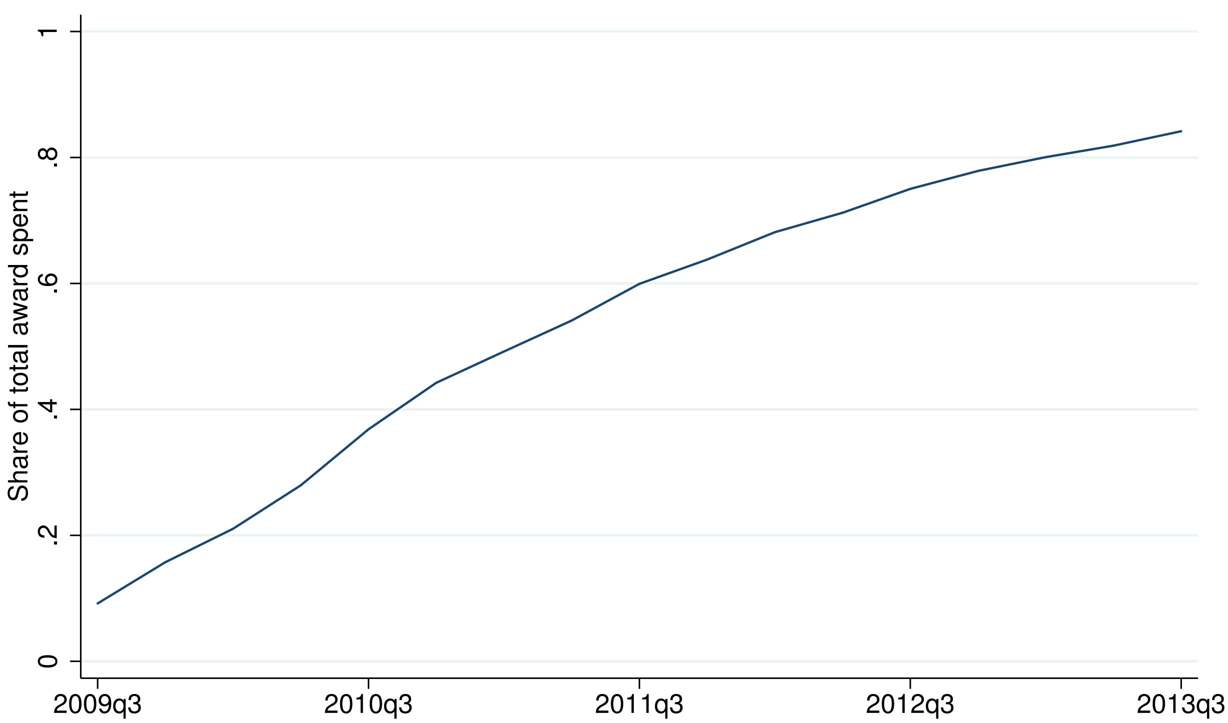


Figure 2: Status adjusted awards per capita in 2011q3

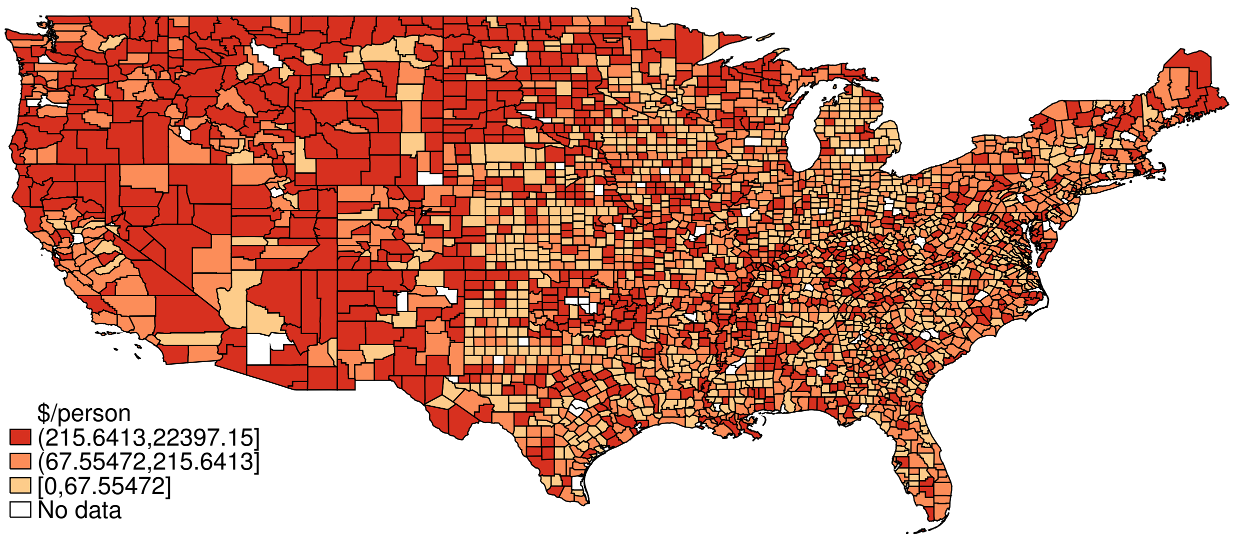
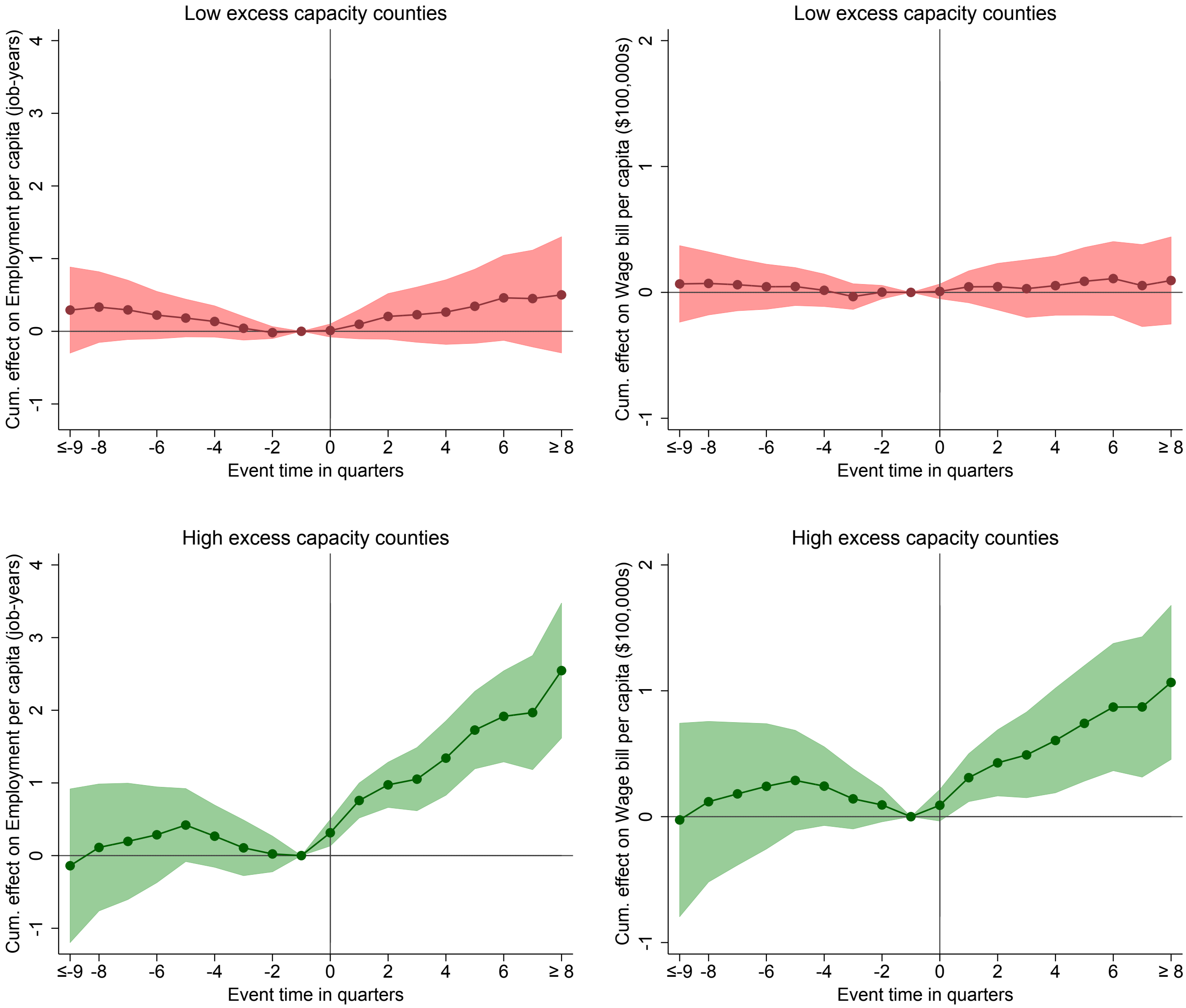
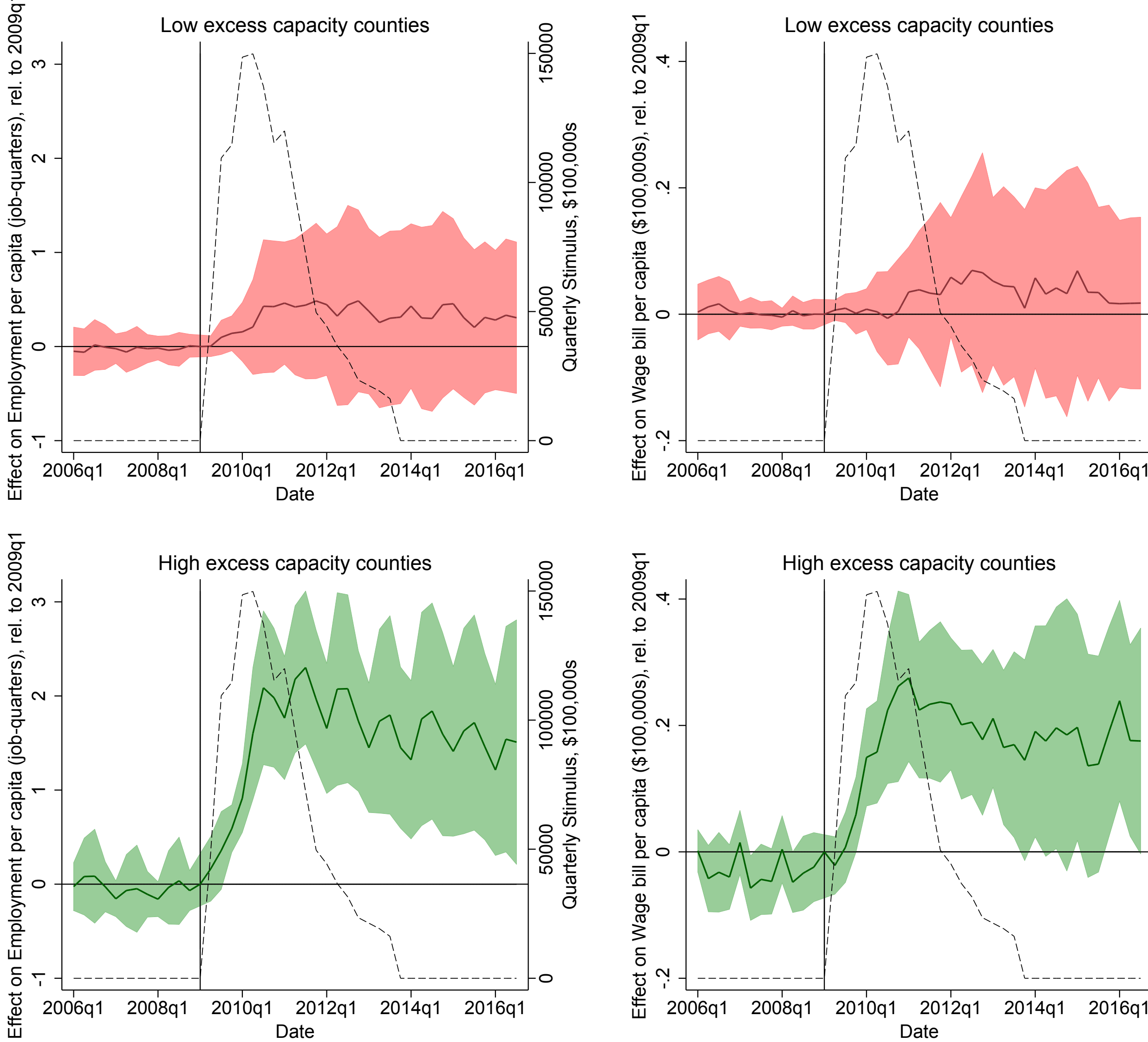


Figure 3: Cumulative response of employment and wages to stimulus, split sample



Notes: Figures show cumulative effects from regressions of own-county employment per capita and own-county wage bill per capita on own-county aggregate ARRA-stimulus, with quarterly lags and leads of treatment. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. Coefficients on lags and leads are (separately) summed cumulatively from event-date -1, where the effect is normalized to 0. The sums of lags include the contemporaneous effect at event-date 0. The vertical reference line indicates 2009q1. The colored line indicates the summed coefficients, while the shaded area is the associated 95% confidence interval. Employment estimates are annualized such that coefficients should be understood as effects on "job-years". Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. Standard errors are clustered at the state level.

Figure 4: Time-based effects of stimulus, using time fixed effects interacted with total award per capita, split sample



Notes: Figures show coefficients from regressions of own-county employment per capita and own-county wage bill per capita on own-county aggregate ARRA-stimulus fully interacted with quarterly time dummies. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. Coefficients are interpreted with reference to 2009q1, the omitted time-dummy. The vertical reference line indicates 2009q1. The colored line indicates the coefficients on the stimulus–time-dummy interaction, while the shaded area is the associated 95% confidence interval. The dashed line indicates the total flow of stimulus awards over time, across all counties. Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. Standard errors are clustered at the state level.

Figure 3: Effects of stimulus, by excess capacity

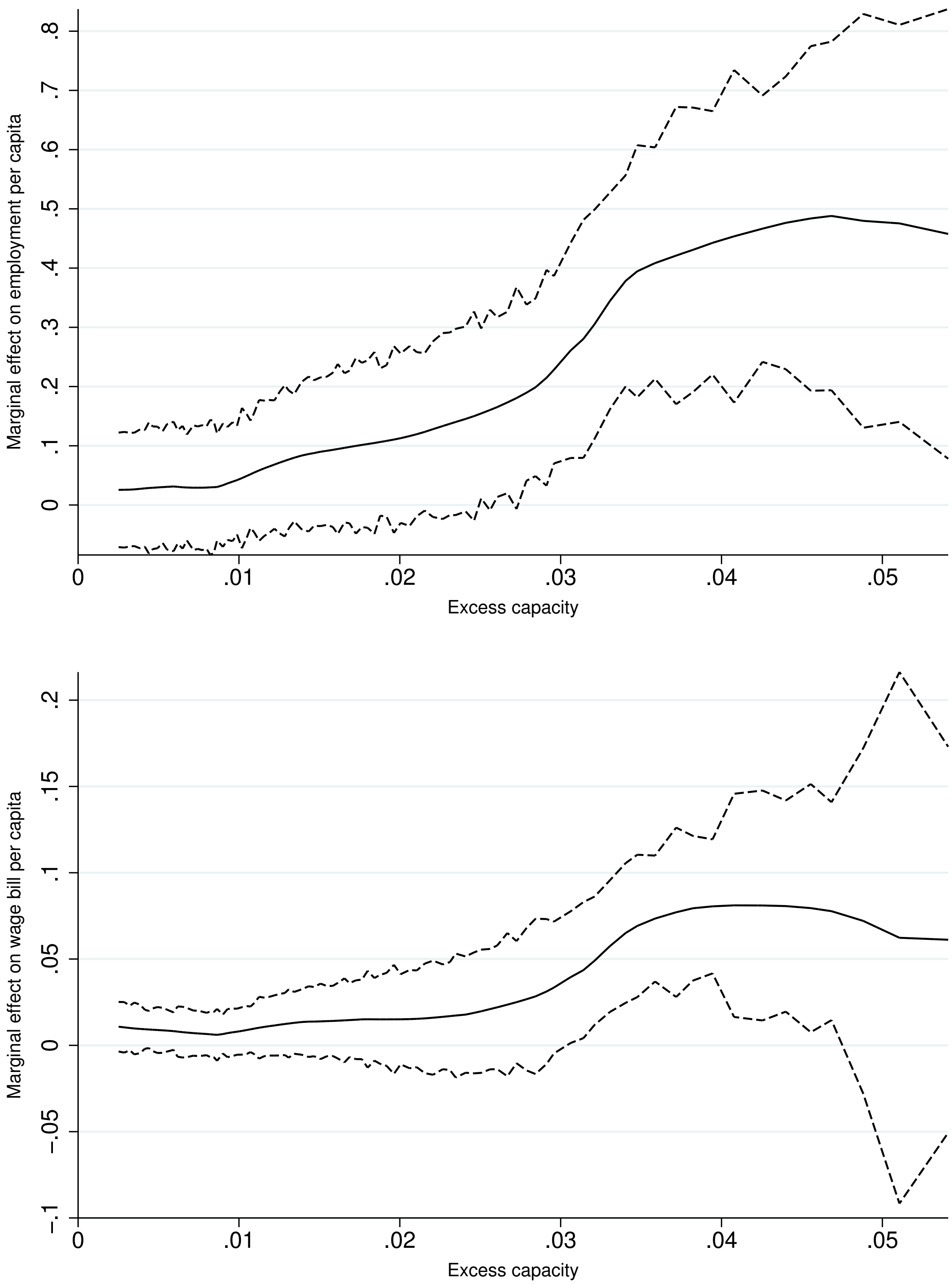


Table 1: Main table

	Employment per capita				Wage bill per capita			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Time-aggregated lags								
Two quarters	0.569***	0.678***	0.414***	0.463***	0.194**	0.196***	0.201**	0.159**
	(0.139)	(0.123)	(0.153)	(0.135)	(0.081)	(0.061)	(0.095)	(0.072)
Four quarters	0.782***	0.987***	0.520**	0.649***	0.305**	0.317***	0.303**	0.251**
	(0.205)	(0.171)	(0.229)	(0.195)	(0.123)	(0.093)	(0.141)	(0.102)
Six quarters	1.252***	1.563***	0.796**	1.012***	0.501***	0.550***	0.416**	0.370***
	(0.312)	(0.250)	(0.318)	(0.254)	(0.175)	(0.115)	(0.196)	(0.128)
Eight quarters	1.629***	2.062***	0.892*	1.211***	0.697**	0.738***	0.489*	0.416**
	(0.505)	(0.390)	(0.466)	(0.349)	(0.278)	(0.170)	(0.277)	(0.161)
Time-aggregated leads								
Eight quarters	−0.253	−0.618	0.319	−0.003	−0.312	−0.569	−0.008	−0.136
	(0.517)	(0.470)	(0.385)	(0.301)	(0.413)	(0.362)	(0.239)	(0.193)
Common time FE	Y	Y			Y	Y		
State X time FE			Y	Y			Y	Y
Controls		Y		Y		Y		Y

Notes: Estimates are of own-county employment and wage bill on own-county stimulus expenditures. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. The outcome variable is employment per capita in the four columns on the left and wage bill per capita in the four columns on the right. Regressions are at the quarterly level, but employment estimates are annualized such that coefficients should be understood as effects on "job-years". Each column shows sums of coefficients from a single regression. The rows under "time-aggregated lags" show the sum of contemporaneous results plus subsequent lags. The row under "time-aggregated leads" shows the sum of coefficients from eight quarter leads. The controls are Bartik predicted employment to population ratio, Bartik predicted wage bill, and demographic controls. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. All specifications include county-level fixed effects. Standard errors are clustered at the state level.

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

Table 2: Excess capacity, split samples

	Employment per capita			Wage bill per capita		
	All counties	Low excess	High excess	All counties	Low excess	High excess
Time-aggregated lags						
Two quarters	0.463***	0.205	0.974***	0.159**	0.046	0.428***
	(0.135)	(0.156)	(0.155)	(0.072)	(0.092)	(0.131)
Four quarters	0.649***	0.264	1.341***	0.251**	0.054	0.605***
	(0.195)	(0.221)	(0.254)	(0.102)	(0.117)	(0.207)
Six quarters	1.012***	0.460	1.916***	0.370***	0.110	0.871***
	(0.254)	(0.291)	(0.311)	(0.128)	(0.146)	(0.251)
Eight quarters	1.211***	0.502	2.546***	0.416**	0.094	1.066***
	(0.349)	(0.397)	(0.461)	(0.161)	(0.172)	(0.304)
Time-aggregated leads						
Eight quarters	−0.003	0.292	−0.139	−0.136	0.067	−0.026
	(0.301)	(0.294)	(0.526)	(0.193)	(0.151)	(0.382)

Notes: Estimates are of own-county employment and wage bill on own-county stimulus expenditures. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. The outcome variable is employment per capita in the three columns on the left and wage bill per capita in the three columns on the right. Regressions are at the quarterly level, but employment estimates are annualized such that coefficients should be understood as effects on "job-years". Each column shows sums of coefficients from a single regression. The rows under "time-aggregated lags" show the sum of contemporaneous results plus subsequent lags. The row under "time-aggregated leads" shows the sum of coefficients from eight quarter leads. Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. Standard errors are clustered at the state level.

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

Table 3: Pre-trend validation tests, split sample

	No controls		Controls	
	Low excess	High excess	Low excess	High excess
Change in employment				
2006q1–2007q1	0.033	0.018	0.003	0.056
	(0.141)	(0.129)	(0.191)	(0.143)
2006q1–2009q1	0.085	0.212	0.033	0.211
	(0.142)	(0.175)	(0.184)	(0.180)
Change in wage bill				
2006q1–2007q1	−0.002	0.068	−0.006	0.065*
	(0.020)	(0.048)	(0.028)	(0.037)
2006q1–2009q1	0.003	0.007	−0.002	0.006
	(0.022)	(0.030)	(0.031)	(0.026)
Pre-ARRA GR severity				
Employment	−0.052	−0.015	−0.039	0.004
	(0.068)	(0.146)	(0.067)	(0.139)
Predicted-employment	−0.049***	−0.115	−0.022**	0.031
	(0.011)	(0.087)	(0.009)	(0.075)

Notes: Estimates are of a variety of trends in pre-ARRA outcome variables on own-county total stimulus expenditure. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. Outcome variables are listed along the rows. The first four rows are changes in employment per capita or wage bill per capita between 2006q1 and 2007q1, or 2006q1 and 2009q1. The fifth and sixth rows are measures of the severity of the Great Recession in a given county before the ARRA was enacted – they are the largest dip in employment per capita and Bartik predicted employment per capita from any quarter in 2006q1–2007q4 to any quarter in 2008q1–2009q1, comparing only between same quarters of the year. Each row per super-column ("No controls" vs "Controls") is a single regression, with "Low excess" and "High excess" specifications estimated simultaneously; each entry is interpretable as coming from a separate (split-sample) regression. Where controls are indicated, regressions include Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. All regressions control for state fixed effects. Standard errors are clustered at the state level.

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

Table 4: Nonlinear effects of the extent of stimulus, split sample

	Employment per capita			Wage bill per capita		
	All counties	Low excess	High excess	All counties	Low excess	High excess
Linear specification	0.281***	0.098	0.805***	0.109*	0.027	0.392**
	(0.092)	(0.095)	(0.201)	(0.055)	(0.062)	(0.170)
Quadratic specification						
Linear term	0.497***	0.180	1.372***	0.188*	0.037	0.679**
	(0.159)	(0.178)	(0.282)	(0.101)	(0.119)	(0.267)
Quadratic term	−15.360	−5.203	−71.573***	−5.612	−0.592	−36.180*
	(9.882)	(8.520)	(25.291)	(4.369)	(4.078)	(20.577)
Fitted effects at:						
25^{th} percentile of stimulus	0.002***	0.001	0.006***	0.810*	0.158	2.916**
	(0.001)	(0.001)	(0.001)	(0.436)	(0.510)	(1.145)
50^{th} percentile of stimulus	0.010***	0.004	0.027***	3.723*	0.727	13.355**
	(0.003)	(0.004)	(0.006)	(2.000)	(2.343)	(5.237)
75^{th} percentile of stimulus	0.025***	0.009	0.069***	9.557*	1.873	34.019**
	(0.008)	(0.009)	(0.014)	(5.120)	(6.009)	(13.302)

Notes: Estimates are of own-county employment and wage bill on own-county stimulus expenditures. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. The outcome variable is employment per capita in the three columns on the left and wage bill per capita in the three columns on the right. Regressions are at the quarterly level, but employment estimates are annualized such that coefficients should be understood as effects on "job-years". The first row reports coefficients on stimulus expenditures in a model with only a linear term for stimulus expenditures. The second and third rows report linear and quadratic coefficients respectively from a single regression with a quadratic polynomial in stimulus expenditures. The "fitted effects" in the fourth to sixth rows present fitted values of the relevant outcome using only the stimulus and stimulus-squared coefficients from the quadratic specification. These are fitted values for counties at the 25th, 50th and 75th percentiles of stimulus as indicated. Fitted effects for employment per capita and wage bill per capita are scaled by 100 and 100,000 respectively, in order to allow interpretation in percentage points and Dollars. Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. Standard errors are clustered at the state level.

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

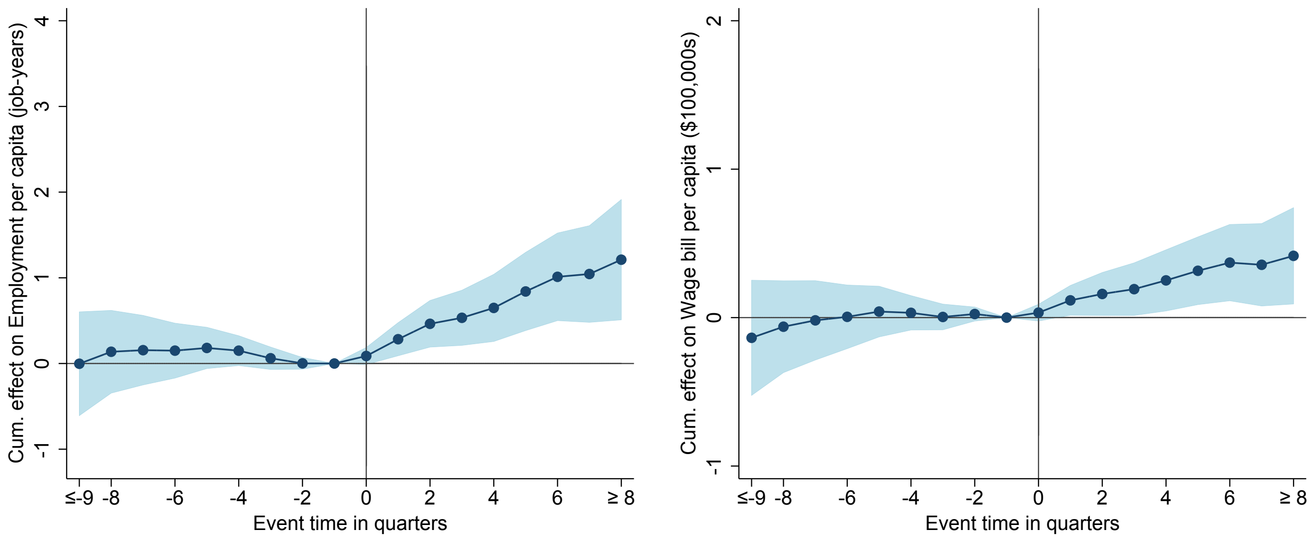
Table 5: Effects of stimulus by industrial sector, split sample

	Employment per capita			Wage bill per capita		
	All counties	Low excess	High excess	All counties	Low excess	High excess
Overall	1.211***	0.502	2.546***	0.416**	0.094	1.066***
	(0.349)	(0.397)	(0.461)	(0.161)	(0.172)	(0.304)
Public sector	0.881***	0.810***	0.784***	0.259***	0.247***	0.254***
	(0.183)	(0.229)	(0.195)	(0.055)	(0.064)	(0.083)
Private sector	−0.107	−0.800*	1.524***	−0.090	−0.437**	0.641**
	(0.366)	(0.445)	(0.379)	(0.181)	(0.216)	(0.277)
Tradables	−0.204	−0.276	−0.044	−0.130	−0.178**	−0.053
	(0.168)	(0.168)	(0.269)	(0.094)	(0.082)	(0.168)
Non-tradables	0.115	−0.271	1.251***	0.029	−0.158	0.496**
	(0.208)	(0.336)	(0.333)	(0.091)	(0.168)	(0.192)
Services	0.179	−0.261	1.159***	0.062	−0.156	0.467**
	(0.199)	(0.306)	(0.269)	(0.094)	(0.155)	(0.198)
Goods	−0.034	−0.083	0.094	−0.043	−0.071	0.020
	(0.159)	(0.160)	(0.220)	(0.099)	(0.086)	(0.145)
Construction	0.090	0.038	0.273*	0.014	−0.014	0.125
	(0.055)	(0.056)	(0.143)	(0.031)	(0.031)	(0.078)

Notes: Estimates are of own-county employment and wage bill on own-county stimulus expenditures. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. The outcome variable is employment per capita in the three columns on the left and wage bill per capita in the three columns on the right. Each estimate is the sum of a contemporaneous effect and eight quarterly lags of stimulus. The full specification includes 8 quarterly leads of stimulus which are not reported. Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Low excess capacity is below and high excess capacity is above the 50th percentile of county excess capacity. Industries come from NAICS classifications in the QCEW. Standard errors are clustered at the state level.

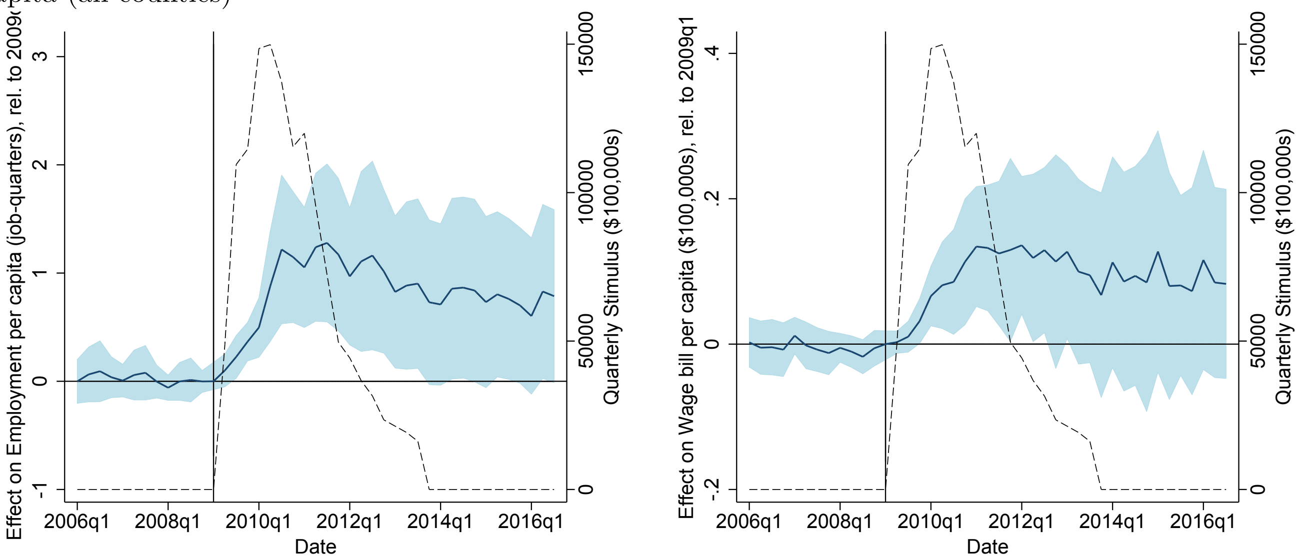
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A1: Cumulative response of employment and wages to stimulus, overall



Notes: Figures show cumulative effects from regressions of own-county employment per capita and own-county wage bill per capita on quarterly leads and lags of own-county aggregate ARRA-stimulus. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. Coefficients on lags and leads are (separately) summed cumulatively from event-date -1, where the effect is normalized to 0. The sums of lags include the contemporaneous effect at event-date 0. The vertical reference line indicates 2009q1. The colored line indicates the summed coefficients, while the shaded area is the associated 95% confidence interval. Employment estimates are annualized such that coefficients should be understood as effects on "job-years". Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend.

Figure A2: Time-based effects of stimulus, using time fixed effects interacted with total award per capita (all counties)



Notes: Figures show coefficients from regressions of own-county employment per capita and own-county wage bill per capita on own-county aggregate ARRA-stimulus fully interacted with quarterly time dummies. Stimulus expenditures and the wage bill are measured in $100,000 per person. County population is the number of residents aged 15–64. The sample is from 2006Q1 to 2016Q3. Employment and wage bill data come from the QCEW. Timing of stimulus expenditures is adjusted from stimulus recipient reports from www.recovery.gov. Coefficients are interpreted with reference to 2009q1, the omitted time-dummy. The vertical reference line indicates 2009q1. The colored line indicates the coefficients on the stimulus–time-dummy interaction, while the shaded area is the associated 95% confidence interval. The dashed line indicates the total flow of stimulus awards over time, across all counties. Regressions control for Bartik predicted employment to population ratio, Bartik predicted wage bill, demographic controls, state-by-time fixed effects and county fixed effects. Bartik predictions are based upon county-level employment and wage bill averages over 2006-2007 at the three-digit NAICS level. Demographic controls are Census 2000 estimates of percents black, hispanic, urban, and under poverty, as well as median income and 2006 average home purchase loans and 2006 total HMDA loans per capita. All demographic controls are interacted with a time trend. Standard errors are clustered at the state level.