Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

A Cup Runneth Over:
Fiscal Policy Spillovers from the 2009 Recovery Act

Bill Dupor
and
Peter B. McCrory

Working Paper 2014-029D
http://research.stlouisfed.org/wp/2014/2014-029.pdf

October 2014
Revised April 2016

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

A Cup Runneth Over: Fiscal Policy Spillovers from the 2009 Recovery Act*

Bill Dupor† and Peter B. McCrory‡

April 1, 2016

Abstract

This paper studies the effects of interregional spillovers from the government spending component of the American Recovery and Reinvestment Act of 2009 (the Recovery Act). Using cross-county Census Journey to Work commuting data, we cluster U.S. counties into local labor markets, each of which we further partition into two subregions. We then compare differential labor market outcomes and Recovery Act spending at the regional and subregional levels using instrumental variables. Our instrument is the sum of spending by federal agencies not instructed to allocate Recovery Act funds according to the severity of local downturns. Among pairs of subregions, we find evidence of fiscal policy spillovers. According to our benchmark specification, \$1 of Recovery Act spending in a subregion increases its own wage bill by \$0.64 and increases the wage bill in its neighboring subregion by \$0.50 during the first two years following the act's passage. We find similar spillover effects when we replace the wage bill with employment as our measure of economic activity. The spillover effect occurs in the service sector, whereas the direct effect occurs in both the services and goods producing sector. Also, we estimate cross-sectional regressions at various levels of aggregation. The estimated effect of stimulus spending increases with the level of aggregation, as greater aggregation subsumes geographic spillovers into the own-region effect of spending.

Keywords: fiscal policy, spillovers, the American Recovery and Reinvestment Act.
JEL Codes: E52, E62.

*The authors thank Tim Conley and Ana Maria Santacreu for useful conversations and also the editor and three referees for valuable comments and suggestions. The authors would also like to thank seminar and conference participants at the Federal Reserve Bank of St. Louis, the Society for Economic Dynamics meeting and the Monetary Policy in a Global Setting conference. A repository containing government documents, data sources, a bibliography and other relevant information pertaining to the Recovery Act is available at billdupor.weebly.com. The analysis and conclusions set forth do not reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Federal Reserve Bank of St. Louis, william.d.dupor@stls.frb.org, billdupor@gmail.com.
‡University of California, Berkeley, pbmccrory@berkeley.edu, peter.mcccrory@gmail.com.

1 Introduction

In response to the 2007-09 recession, the U.S. government enacted the American Recovery and Reinvestment Act of 2009 (hereafter, Recovery Act). The Recovery Act was the largest countercyclical fiscal intervention in the U.S. since FDR's New Deal. The law's total budget impact was $840 billion. Drautzburg and Uhlig (2013) report that $350 billion of this amount constituted government purchases of goods and services.[1]

The act was a massive commitment from the federal government to many sectors of the economy, including highway infrastructure, energy and education. For example, the U.S. Department of Education distributed $94 billion in Recovery Act spending, which equals nearly $2,000 per elementary/secondary public school student.

In this paper, we estimate the extent to which the Recovery Act increased local economic activity as well as how this impact propagated itself geographically. Our starting point is the observation that roughly 34% of workers in a typical county are employed outside their county of residence.[2]

As an example, of the 1 million workers residing in Brooklyn (Kings County, NY), 50% are employed in a different county. Seventy-four percent of these commuters from Brooklyn work in nearby Manhattan (New York County, NY), representing approximately one of every five workers in Manhattan. Not only do a sizable number of Brooklyn residents work in Manhattan, these commuters comprise a substantial portion of all Manhattan workers. The degree of economic interdependence between these New York City boroughs is high.

Suppose the federal government increased its purchases in Manhattan. Since many Brooklyn residents earn their income in Manhattan, presumably these commuters would spend a significant part of their income in their home county. Government purchases in Manhattan could in turn increase consumer purchases, and other measures of economic activity, in Brooklyn. This potential for cross-county economic interdependence motivates our analysis and our methodological approach.

Using county-level job commuting data, we organize the U.S. into 1293 distinct local labor markets, or regions. We then partition each of these regions into two subregions: a large county subregion and a satellite subregion; the latter is the aggregation of all of the remaining counties within the region. We then ask: how does government spending in one subregion affect its own economic activity as well as the economic activity of its partner subregion?

We measure economic activity in each subregion by its employment level and wage bill.[3] We measure counter-cyclical government spending using quarterly reports filed by over 570,000 recip-

[1] The remainder reflected direct entitlement payments to individuals and tax cuts. Some existing studies, and interpretations of those studies, substantially understate the amount of government purchases the act generated. Bureau of Economic Analysis (2013) mislabels over $100 billion as transfer payments that more accurately should be treated as state and local government consumption and investment. Cogan and Taylor (2012) state that in 2009 and 2010 federal government purchases resulting from the act totaled only $30 billion; however, that number does not include state and local government consumption and investment.

[2] Authors' calculation using county-to-county commuting flows reported in the American Community Survey.

[3] Unfortunately, data on gross domestic product and its components are not available at the county level.

ients (businesses, nonfederal government agencies and nonprofit organizations) of Recovery Act funds. These reports provide zip-code-level detail on spending, allowing us to execute a highly disaggregated analysis.

We have four main findings. First, we find that Recovery Act spending in a geographic area increased wage payments and employment in that area (relative to a no stimulus counterfactual); moreover, spending in one area *spilled over* to nearby communities and similarly increased wage payments and employment there. In the first two year's following the act's passage, \$1 of Recovery Act spending in one part of a labor market region increased that part's own wage bill by \$0.64 (SE = 0.22) and increased the wage bill in the rest of the region by \$0.50 (SE = 0.07).

Second, we find similar effects when we replace the wage bill with the level of employment as our measure of economic activity: over the same horizon, \$1 million of stimulus in one part of a local labor market increased employment there by 10.3 (SE = 3.8) persons and increased employment in the rest of the region by 8.5 (SE = 2.8) persons.

Third, although we find strong evidence of substantial spillovers *within* regional markets, we find no convincing evidence for similar spillovers *between* such markets. Again grouping counties on the basis of commuter flows, we estimate the causal impact of spending on regions defined at varying levels of county aggregation. The estimated labor market effects of Recovery Act spending increase with the level of aggregation, as greater aggregation subsumes geographic spillovers into the own-region effect of spending. Beyond a certain point of aggregation, our estimates become relatively stable. We perform a randomized placebo test and show that between-county spillovers disappear when we shutdown the commuting linkages between counties located in the same region.

Fourth, we present a sectoral decomposition of the direct and spillover effects of spending and analyze the dynamic impact of the spending on labor market variables.

Overall, our results lend support to the theory of the "government spending multiplier." That is, government spending not only combats recessions through directly increasing economic activity and hours worked, but also through spillover effects as additional income in workers' hands is spent.

We use instrumental variables to address potential endogeneity in the allocation of Recovery Act spending. Our instrument is the sum of spending from a number of Recovery Act programs that we find did not allocate funds to more economically distressed, or alternatively strong, areas. The allocation of funds through these programs is plausibly conditionally uncorrelated with the business cycle conditions in a particular local labor market. We carry out this task by analyzing the act, federal codes and regulations cited by the act, and implementation guidances written by the agencies tasked with allocating the funds. Examples of these components include the Energy Efficiency and Renewable Energy program (Department of Energy), the Capital Transit Assistance Program (Federal Transit Administration), the Special Education Fund (Department of Education), and the Public Building Fund (General Services Administration).

We then instrument overall Recovery Act spending with the subset of stimulus spending that is conditionally uncorrelated with the local area business cycle. Then, we compare differences in labor market outcomes with differences in predicted Recovery Act spending across observations to estimate the direct and spillover causal effects.

Our paper relates to two lines of research. First, our general methodology follows other studies that use cross-sectional instrumental variable techniques to estimate the economic impact of spending at the subnational level. These include Shoag (2012), who studies the effect of unanticipated capital gains to government pension funds, and Clemens and Miran (2012), who use differences in state balanced budget requirements to identify the effects of fiscal stabilization policy. A few papers also apply this general methodology to the Recovery Act episode. These include Chodorow-Reich et al. (2012) and Wilson (2012).

Second, our paper is related to research on fiscal policy spillovers. Carlino and Inman (2013) study a panel of U.S. states and show that an exogenous increase in one state's deficit can generate increased employment in neighboring states. Beetsma and Giuliodori (2011) apply short-run restrictions to identify the effects of government spending shocks in the European Union using a structural vector autoregression. They find that an exogenous spending shock in a large European Union country increases output in other union member countries.

The closest paper to ours with respect to studying spillovers is Suarez Serráto and Wingender (2014). Using a different instrument and a different time period than our study, they estimate the direct and spillover effects of government spending using cross-sectional methods. They find positive spillovers between counties that are geographically close to one another.

2 Empirical Analysis

2.1 The Data

The local labor market

We begin by defining a local labor market as a set of counties with an interdependent economic structure. We use cross-county commuting patterns to identify both economic spatial dependence among counties and spatial economic independence across local labor markets. We construct local labor markets with two conditions in mind: within-market dependence and across-market independence. Hereafter, we refer to local labor markets as regions or regional markets, interchangeably.

Following a methodology similar to that of Tolbert and Sizer (1996), we use an agglomerative hierarchical clustering technique to identify such independent regional markets.[4] This approach makes the implicit assumption that commuting patterns are good proxies for economic interdependence.

First, we construct a single nationwide pairwise flow matrix from county-to-county commuting data (the 2000 Journey to Work survey) that measures relative economic distance between any two

[4]These are referred to as commuting zones in their paper. For examples of applications of the commuting zone approach, see Autor, Dorn and Hanson (2013) and Chetty et al. (2014).

counties. We define the economic distance between county i and county j as

$$D_{i,j} = D_{j,i} = 1 - \frac{(C_{i,j}) + (C_{j,i})}{min(LF_i, LF_j)} \tag{2.1}$$

where $C_{i,j}$ indicates the number of commuters from counties i to j. LF_i refers to the employed resident labor force in county i.[5] With this measure of economic distance, we use the average linkage algorithm to map the pairwise flow matrix to x clusters of counties (regional markets)—with x dependent on a threshold parameter, γ, that indicates the average distance between clusters. By lowering the average distance between regions, we increase the number of regions identified. Likewise, by increasing the average distance between regions, we can group together increasingly disparate, less-economically interdependent counties.

The choice of γ balances two possible costs: (i) setting γ too low, in which case one might fail to group together counties that are, in fact, economically interdependent or (ii) setting γ too high, in which case one might incorrectly conclude that some set of relatively isolated counties together form a single regional market.[6]

Because we are primarily interested in identifying spillovers, changing the γ parameter allows us to identify sets of counties with stronger connections between them than the official USDA delineation of the country into commuting zones in 2000. Our baseline set of 1293 regional markets is a more granular partition than the USDA set, which partitions the country into 709 regional markets. By choosing the more granular partition instead of the official USDA delineation we improve the statistical power of our analysis by increasing our sample size while simultaneously increasing the average strength of the linkage among counties within a given regional market.

Partitioning the local labor market

Since our purpose is to identify the short-run effect of an injection of spending into a local economy and the geographic spillover of that effect, a more systematic discussion of the regional markets is in order. Note that when referring to a particular set of identified regions, we use LM_J, where J indicates the total number of regions in the set. Our baseline choice is LM_{1293}.[7] Unless otherwise noted, results are from this baseline partition.

[5] Our measure of distance differs slightly from that of Tolbert and Sizer (1996) in that we use the *employed* resident labor force, whereas they use the entire resident labor force. We also use a nationwide flow matrix rather than overlapping regional matrices to identify clusters.

[6] Tolbert and Sizer (1996) delineate regions with γ set to 0.98. Using the 2000 data, this parameter results in a slightly more agglomerated mapping than the official U.S. Department of Agriculture case. This arises primarily because we use a nationwide flow matrix rather than overlapping regional matrices. In our benchmark designation, we set $\gamma = 0.93$. This value was chosen because it produced a classification at which within local labor market spillovers become discernible.

[7] In our estimates, J does not always equal the sample size because we restrict our analysis to markets with more than 25,000 residents. This and the additional condition that more than one county be present in a market further restrict the sample size for the spillover analysis.

First, *we subdivide each region into a pair of subregions.*[8] Each pair consists of the largest county and the sum of the remaining counties in the region. We denote variables pertaining to the largest county in a particular region with the subscript $s = 1$ and refer to them collectively as the large county subregions.

The remaining satellite counties in the region—that is, excluding the largest county—constitute the second level of observation. We use the subscript $s = -1$ to refer to this set within a region. The pair (j, s) refers to the subregion s from regional market j and $(j, -s)$ refers to its adjacent subregion.

Whenever a subregion is comprised of multiple counties rather than a single county, we construct variables for that subregion in the following way. We take the level values from the constituent counties and combine them as befitting the form needed. For instance, for N counties in subregion (j, s), the natural log of the subregion's population is given by $lnpop_{j,s} = \log\left(\sum_{n=1}^{N} Population_{n,s}\right)$. *Mutatis mutandis* are other requisite variables constructed.

As an illustration of how this clustering approach operates, Figure 1 presents the regional market partition of Pennsylvania. We delineate regions by color so that no region shares a color with an adjacent region. That is, any contiguous mapping of counties of the same color corresponds to a *single* regional market. To further indicate subregions, we color the large county subregion with a darker color tone. For example, in the bottom-right quadrant of the map we identify Philadelphia County, which contains the eponymous city of Philadelphia, as the largest county in its regional market. As the large county subregion in the four-county regional market, Philadelphia County is colored dark blue. The corresponding satellite subregion, consisting of the remaining three counties, is colored light blue.

Counties that are not contained within a regional market with any other Pennsylvania county are colored gray. Either these counties comprise single-county regional markets or they are, via commuting linkages, more tightly connected to counties outside the state.[9] The LM_{1293} set partitions Pennsylvania into 22 regional markets with at least two counties in the state. To show how these regional markets reflect the distribution of population across the state, we indicate the 57 cities in Pennsylvania with black ovals that are proportional to the city population in 2010.

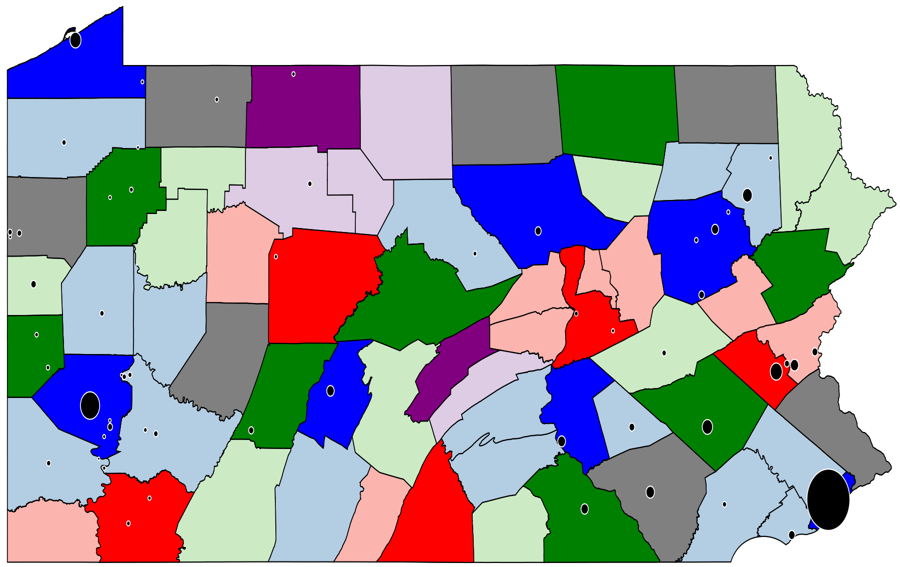
We assess how LM_{1293} partitions the nation into regional markets in Table 1. First, we group regions by the number of counties contained within them. These groups are listed in column (1). Next, for every region identified, we construct the following ratio: $ratio_{LM_{1293}} = \frac{Pop_{j,1}}{Pop_{j,-1}}$, which is the population of the large county subregion relative to the population of the satellite subregion.

For each grouping by number of counties, we report the average of ratios in column (2). Column (3) tabulates the number of regions used to construct the ratio averages. For example, there are

[8]This subdivision is not possible for 323 regions consisting of a single county. We drop these from our subregional analysis.

[9]This illustrates one additional benefit of our approach: the identified regional markets are not confined by state borders.

Figure 1: Local labor markets (regions) in Pennsylvania



Notes: Any contiguous mapping of a single color represents a specific regional market from LM_{1293}. The large county subregion is represented by the darker tone and the satellite subregion by the lighter tone. Black ovals indicate the 57 cities in Pennsylvania. Oval sizes are proportional to city population in 2010.

443 regions that are comprised of three counties; for this group, the average ratio between the population of the largest county and the population of the two remaining counties is 2.32. Column (4) provides the total population for each regional market grouping. Note that we exclude Alaskan regions where commuting patterns and local market conditions are likely to differ substantially from those across the rest of the nation.

For the majority of regions in LM_{1293} and for a sizable portion of the population, the largest county population, on average, is larger than the rest of the counties combined. This implies that many regional markets have a distinctively unimodal structure when disaggregated by county.[10]

Outcome variables

We explore two alternative outcome variables: the wage bill and employment. The wage bill and employment data are from the Quarterly Census of Employment and Wages (QCEW), which covers approximately 98% of U.S. jobs. We focus first on the wage bill response to government spending. Total wages received by employees in a given quarter also include "bonuses, stock options, severance pay, profit distributions, cash value of meals and lodging, tips and other gratuities, and, in some

[10] The ratio of the large county subregion to the satellite subregion is inversely correlated with the size of the regional market, implying that the unimodal structure does not hold among the most populous markets.

Table 1: Population distributions of regional markets by number of constituent counties from LM_{1293}

Number of Counties (1)	Ratio (2)	Number of LMs (3)	Population (millions) (4)
1	-	323	14
2	4.43	443	52
3	2.32	288	72
4	2.12	128	55
5	1.29	56	56
6	1.25	18	23
7	1.09	5	5
8	2.07	7	24
10	0.44	2	6
All LMs	3.19	1,270	308

Notes: Population distributions from the aggregation identifying 1293 regions (labor markets [LMs] in table). The statistics above exclude Alaska, where commuting patterns and the economics of regional markets are likely to differ substantially from those across the rest of the nation.

States, employer contributions to certain deferred compensation plans such as 401(k) plans."[11] These data are observed at the county level and are mapped to the regions identified above. Table 2 contains summary statistics for the variables used in our analysis.

More formally, the outcome variable is the accumulated change in the per capita wage bill in the two years following the passage of the Recovery Act, relative to a base-period of 2008:Q4. That is,

$$\Delta\text{Wage-Bill}_{j,s} = \frac{1}{Pop_{j,s} + Pop_{j,-s}} \sum_{k \in K} (P_{j,s,k} - P_{j,s,2008Q4}) \tag{2.2}$$

where j indicates a particular region, s a particular subregion, k indicates the quarter, $K = \{2009 : Q1, ..., 2010 : Q4\}$, and $P_{j,s,k}$ indicates the total wages received by employees in given quarter.

Treatment variables ($ARRA_{j,s}$)

Define $A\bar{R}RA_{j,s}$ as the cumulative value of Recovery Act dollars through 2010Q4 spent by organizations in subregion s in region j. These amounts are constructed from quarterly reports filed by all recipients of contracts, grants, and loans. The data were downloaded from Recovery.Gov.[12]

[11] See: http://www.bls.gov/cew/cewfaq.htm#Q15 for a description of the QCEW data. These wages do not include other forms of worker compensation, such as medical insurance.

[12] We use "spending" and "aid" interchangeably throughout the paper to refer to the dollar amount of Recovery Act funds spent by prime recipients, net of expenditures made by sub-recipients and payments to vendors and sub-vendors. In Table A.6 we consider the treatment to be the the value of the award made to primary recipients and sub-recipients net of payments to vendors. Our results are qualitatively similar when using this alternative treatment

Table 2: Summary statistics for regional market variables used in estimating spatial spillovers, LM_{1293}

	Mean	SD	10th Percentile	90th Percentile
Large County Subregion				
Δ Job-years p.c., (2008Q4-2010Q4)	-0.02	0.02	-0.04	-0.00
Δ Wage bill p.c., (2008Q4-2010Q4)	-1,579.98	1,385.65	-2,890.58	-343.06
Recovery Act Expenditure p.c.	283.77	305.42	89.03	512.72
Adjacent ARRA Expenditure p.c.	124.35	141.40	24.01	244.80
Composite Instrument Expenditure p.c.	56.85	57.13	13.79	115.62
Adjacent Composite Instrument Expenditure p.c.	30.75	56.14	2.02	61.68
Subregional market income p.c. (3-yr MA)	487.76	519.16	11.21	1,061.33
Log of subregional population	11.38	1.18	10.04	13.11
Subregional manufacturing share	0.14	0.09	0.04	0.27
Δ Unemployment Rate, (Jan. 2008-Jan. 2009)	0.03	0.02	0.01	0.05
Subregional market job level p.c. (2007Q4)	0.28	0.09	0.16	0.39
Subregional Wage bill level p.c. (2007Q4)	2,525.56	1,052.94	1,324.33	3,885.01
Satellite Subregion				
Δ Job-years p.c., (2008Q4-2010Q4)	-0.01	0.01	-0.02	0.00
Δ Wage bill p.c., (2008Q4-2010Q4)	-626.34	796.94	-1,347.59	-42.43
Recovery Act Expenditure p.c.	124.35	141.40	24.01	244.80
Adjacent ARRA Expenditure p.c.	283.77	305.42	89.03	512.72
Composite Instrument Expenditure p.c.	30.75	56.14	2.02	61.68
Adjacent Composite Instrument Expenditure p.c.	56.85	57.13	13.79	115.62
Subregional market income p.c. (3-yr MA)	196.01	265.52	-39.16	518.98
Log of subregional population	10.69	1.36	9.19	12.50
Subregional manufacturing share	0.15	0.10	0.04	0.29
Δ Unemployment Rate, (Jan. 2008-Jan. 2009)	0.03	0.02	0.01	0.06
Subregional market job level p.c. (2007Q4)	0.11	0.07	0.04	0.20
Subregional Wage bill level p.c. (2007Q4)	1,009.55	788.89	298.77	1,803.67

Notes: The statistics above exclude Alaska and regional markets with fewer than 25,000 residents. SR indicates subregion; ARRA, American Recovery and Reinvestment Act of 2009; FHWA, Federal Highway Administration; p.c., per person in the regional market; SD, standard deviation. All variables above are reported in level or per capita terms. The wage bill and personal income variables are scaled to be per million in the regressions.

Table 3: Components of the Recovery Act used in the construction of the instrument

Federal Department/Agency	Program Title	Amount Authorized (in billions)
Environmental Protection Agency	State and Tribal Assistance Grants	7.2
General Services Administration	Public Building Fund	5.6
General Services Administration	Energy Efficient Federal Motor Vehicle Fleet Procurement	0.3
Department of Education	Special Education Fund	12.2
Department of Energy	Energy Efficiency and Renewable Energy	16.5
Department of Justice	Office of Justice Programs	2.7
Federal Transit Administration	Capital Transit Assistance (Urban and Non-Urban Programs)	6.9
U.S. Army Corps of Engineers	Civil Program Financing Only-Construction	2.1
U.S. Army Corps of Engineers	Civil Program Financing Only-Operation and Maintenance	2.0

In these reports, recipients provide the place of performance zip code, which allows us to map the amount received by prime recipients, subrecipients, vendors, and subvendors to a particular county, net of any portion of the funding reported as spent by a different entity.[13]

We scale the cumulative Recovery Act expenditures to each subregion s by the overall regional market population in j and report the variable in terms of millions of dollars:

$$ARRA_{j,s} = \frac{A\bar{R}RA_{j,s}}{(1e+6) \times (Pop_{j,s} + Pop_{j,-s})} \tag{2.3}$$

Instrument variables ($CompARRA_{j,s}$ **)**

Because policymakers intended to distribute some of the funds to regions most affected by the recession, estimation by least squares might suffer from an endogeneity bias. To ameliorate this bias, we look for components of the Recovery Act for which the allocation of funds was plausibly uncorrelated with the business cycle in a particular local labor market. We identify these components by analyzing the act, federal codes and regulations cited by the act and implementation guidances that were written by the agencies tasked with allocating the funds. We argue that the funds distributed through these Recovery Act programs were neither allocated to more economically distressed regions nor, alternatively, to areas with relatively strong economies.

definition.

[13]For instance, a prime recipient might spend a portion of the award in the county in which it operates while redistributing funding to a subrecipient operating in an adjacent county. Recipients were required to report only payments made to subrecipients, vendors, and subvendors in excess of $25,000. A casual review of the data shows that many payments less than $25,000 were also reported. Though we cannot observe recipient level data for these unreported awards, we do know the total value of unreported awards by recipient. In total, these small awards represent less than 3% of all funding reported in the recipient reports. Finally, vendor expenditures are mapped to a particular market through the zip code location of its headquarters.

This is a straightforward exercise since almost every agency provided at least one detailed plan describing the criteria by which funds would be allocated. Moreover, we choose categories that are representative of the goods and services which the Recovery Act purchased. Most of these can be divided into either infrastructure spending or aid to local and state governments. If the categories that we selected were not representative of the overall Recovery Act spending, then, to the extent that the effects of, say, infrastructure spending are different than that of aid to governments, we would over represent one of the two and in turn bias our results.

We provide evidence for the exogeneity of two components of our instrument here and discuss the other components in the appendix. First, we consider Environmental Protection Agency (EPA) State and Tribal Assistance Grants. The Recovery Act included $7.22 billion for EPA projects. The largest EPA programs were the State Revolving Fund Capitalization Grants to supplement the federal Clean Water State Revolving Fund and the Drinking Water State Revolving Fund, for which the act allocated $4 and $2 billion respectively. Since the capitalization grants were the lion's share of the EPA's entire stake in the Recovery Act, our discussion of the EPA's funding guidelines will be restricted to these programs.

States prepared annual Intended Use Plans to describe how funds would be distributed. An administrative guidance, Environmental Protection Agency (2009), describes several of the criteria that states were to use in their own project selection. These include giving priority to projects that will be "ready to proceed to construction within 12 months of enactment of the Act," and having "not less than 20% of funds go to green projects." There were also "Buy American" requirements for iron, steel and manufactured goods incorporated into projects and Davis-Bacon wage rate restrictions. Nowhere in the guidances that we read or the legislation itself is there mention of states being directed to apply fund to areas hardest hit by the recession.[14] Given the federal guidances, we argue that program administrators–at the state level–would put much greater concern towards putting money where water quality needs were greatest as opposed to attempting to use funds to combat low employment in particular counties within a state.

Second, consider the Department of Justice Office of Justice Programs (OJP). Grants were administered to state and local governments to support activities "to prevent and control crime and to improve the criminal justice system."[15] The program was authorized $2.7 billion. Of this amount, $1.98 billion was issued via formulary Justice Assistance Grants (JAG). Sixty percent of the JAG allocation was awarded to states with the remainder set aside for local governments. The formula dictating allocations is based on population and violent crime statistics. The formula also includes minimum allocation rules to prevent states and localities from receiving disproportionately low funds. The next three largest components of the OJP were for correctional facilities on tribal lands ($225 million), grants to improve the functioning of the criminal justice system ($125.3 million) and rural law enforcement grants to combat crime and drugs ($123.8 million). All three

[14] These documents include Environmental Protection Agency (2009) and Environmental Protection Agency (2011).
[15] See Department of Justice (2009a).

were discretionary grants.

Nowhere in the program's documentation that we examined do we find instructions from the Department of Justice to have localities or states direct grant aid to those areas harder hit by the recession. For example, with respect to the correctional facilities on tribal lands grants, there are a number of restrictions (see Department of Justice (2009b)). A few of these are "Buy American" provisions, Bacon-Davis wage requirements and preference for quick start activities. Serving areas hardest hit by the recession as an instruction to recipients or a criterion for receiving the grant is not among the restrictions. We conclude that the allocations of this component of the act were largely uncorrelated with the degree of economic weakness in the local labor markets that received this aid.

Now, define $CompARRA_{j,s}$ as the cumulative Recovery Act funding through each of the components reported in Table 3 to subregion s in region j, constructed in the same fashion as overall Recovery Act expenditures, $ARRA_{j,s}$.

Conditioning Variables

We estimate our benchmark models with three sets of control variables along with a constant.[16]

- The first set consists of eight Census region dummies.
- The second set includes ten wage bill trend controls, five controlling for own subregion trends and five controlling for wage bill trends in the adjacent subregion. They are per capita wage bills in 2008Q4 and the preceding four quarters.
- The third set of controls pertains to the idiosyncratic economic conditions and structure in each subregion. We include in this set the share of employment in manufacturing, the natural log of population, a 3-year moving average of annual personal income per capita (from 2006 to 2008), and the change in the unemployment rate between January 2008 and January 2009.

3 Estimation and Results

Let $X_{j,s}$ denote the conditioning variables. Let $j = 1, .., N$ denote every regional market from LM_{1293} that contains at least two counties. As explained above, we partition each of these regions into a large county subregion and an adjacent satellite subregion and index large county subregions with $s = 1$ and satellite subregions with $s = -1$.

[16] Note that in the spillover regressions, the numerator for the per capita variables is observed at the subregional level and then scaled by the regional market population.

Our benchmark specification is

$$
\begin{aligned}
\Delta\text{Wage-Bill}_{j,s} &= \psi^D ARRA_{j,s} + \psi^S ARRA_{j,-s} + X_{j,s}\beta + \epsilon_{j,s} \\
ARRA_{j,s} &= \eta^D CompARRA_{j,s} + \eta^S CompARRA_{j,-s} + X_{j,s}\Gamma + v_{j,s} \\
ARRA_{j,-s} &= \eta^{S^*} CompARRA_{j,s} + \eta^{D^*} CompARRA_{j,-s} + X_{j,s}\Gamma^* + v_{j,-s}
\end{aligned} \tag{3.1}
$$

for $j = 1, \ldots, N$ and $s = -1, 1$. Hence, there are N regional markets and $2N$ observations.

Our econometric model parses the direct effect of spending in a sub-region, ψ^D, from the indirect effect of spending coming from the corresponding subregion, ψ^S. Our model imposes a symmetric spillover effect. That is, the spillover response from the large county to its outlying subregion is the same as that from the outlying to the large county subregion. In Section 4.3, we explore our results when we loosen this symmetry restriction, allowing effects to vary by where the fiscal policy intervention occurs within the regional market.

We estimate the model by two-stage least squares (2SLS) with robust standard errors and standard errors clustered by state.[17]

3.1 First Stage Results

To establish the properties, including the strength of the instruments, we examine the first stage results. Note that there are two sets of first stage results: one for each outcome variable, i.e. job years and the wage bill, because each outcome variable includes lags of itself in the corresponding econometric specification.

We present the wage bill first-stage results first. Because there are two endogenous variables, ARRA spending in the own subregion and ARRA spending in the adjacent subregion, we have one table of results for each variable. These are presented in Tables 4 and 5

[17]Each multi-state satellite subregion is mapped to the state in which the bulk of its population resides.

Table 4: First stage least squares estimates of the effect of the composite Recovery Act spending upon direct spending, wage bill model of direct and spillover results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Composite Instrument	1.65***	1.59***	1.54***	1.48***	1.46***
expenditure ($1 Million p.c.)	(0.18)	(0.16)	(0.15)	(0.13)	(0.13)
Adjacent Composite Instrument	0.06	0.06	0.04	0.07	0.05
expenditure ($1 Million p.c.)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
Wage bill level (2007Q4)	0.06***	-0.08	-0.06	-0.09	-0.07
	(0.01)	(0.12)	(0.11)	(0.10)	(0.10)
Added Wage Bill Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Wage Bill Lags	No	No	No	Yes	Yes
Non Labor Controls	No	No	Yes	Yes	Yes
Extra Controls	No	No	No	No	Yes
N	1630	1630	1601	1601	1601
R^2	0.291	0.311	0.311	0.324	0.327
Kleibergen-Paap F-statistic	47.498	54.783	54.476	69.368	66.091

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 5: First stage least squares estimates of the effect of the composite Recovery Act spending upon spillover spending, wage bill model of direct and spillover results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Composite Instrument	0.08	0.10	0.11	0.09	0.08
expenditure ($1 Million p.c.)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)
Adjacent Composite Instrument	1.82***	1.81***	1.74***	1.49***	1.48***
expenditure ($1 Million p.c.)	(0.14)	(0.13)	(0.15)	(0.12)	(0.12)
Wage bill level (2007Q4)	-0.05***	-0.03	0.02	0.03	0.05
	(0.01)	(0.05)	(0.05)	(0.04)	(0.04)
Added Wage Bill Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Wage Bill Lags	No	No	No	Yes	Yes
Non Labor Controls	No	No	Yes	Yes	Yes
Extra Controls	No	No	No	No	Yes
N	1630	1630	1601	1601	1601
R^2	0.258	0.262	0.287	0.346	0.349
Kleibergen-Paap F-statistic	47.498	54.783	54.476	69.368	66.091

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

The own-subregion first-stage results appear in Table 4. In column (1), the only exogenous conditioning variable is the 2007Q4 per capita wage bill in the own-subregion. The coefficient on the composite instrument equals 1.65. Thus, for each dollar of composite spending, there is $1.65 of total Recovery Act spending. The t-statistic for the estimate is 9.17 which gives an indication of a strong instrument. The coefficient on the other instrument, adjacent composite spending is close to zero, with a t-statistic less than 1. This is not surprising since we have no reason to believe that spending in the composite categories in one sub-region should predict overall spending in an adjacent subregion.

In columns (2) through (3), we sequentially add additional conditioning variables: wage bill lags and non-labor controls. None of these lead to a substantial change in the values of the coefficients for either instrument. In our benchmark specification, column (4), which also includes census region dummies and spillover wage bill lags, the coefficient on composite spending equals 1.48 and the coefficient on adjacent composite spending equals 0.07. The final column adds extra conditioning variables: the proportion of employment the tradable sector, the percent change in house prices between 2002Q4 and 2005Q4, and the percent change in house prices between 2005Q4 and 2009Q4. The coefficients on the instruments barely budge.

Table 5 presents the first-stage results from the other endogenous variable in the wage bill equation: Recovery Act spending in adjacent subregions. The structure of the table mimics that of the previous one. Column (1), which conditions only on the 2007Q4 wage bill, shows that a one dollar increase in composite spending in an adjacent subregion leads to $1.82 in stimulus in the adjacent subregion. Second, not surprisingly, there is no statistically or economically significant effect of composite instrument spending on adjacent stimulus spending. Adding additional conditioning variables, i.e. moving from column (1) sequentially to column (5), has a only small effect on the coefficients on the instruments in the regression. In the benchmark specification, the direct and adjacent coefficients are 0.09 and 1.49. The t-statistic on the adjacent composite spending is large, providing evidence that adjacent composite spending is a strong instrument for adjacent Recovery Act spending.

For the job-years specification, there are another set of first-stage results (see Tables A.2 and A.3 in the appendix). These differ from those in Tables 4 and 5 only in that the lagged dependent variables change from wage bill per capita to employment per capita. The impact of this change on the coefficients of interest is very small.

3.2 Wage Bill Estimates

The estimates of (3.1) are given in Table 6. Column (1) presents the estimates when we do not control for regional conditions, labor market characteristics, or wage trend patterns apart from the pre-recession wage bill. The direct effect coefficient, ψ^D, is 0.88 (SE = 0.26). This indicates that an additional $1 in Recovery Act funding in a subregion is associated with an increase in the wage bill of $0.88 in that subregion. The spillover coefficient, ψ^S, equals 0.46 (SE=0.15). Thus, $1 of funding to a subregion is associated with an increased wage bill in the adjacent subregion of $0.46.

Column (2) of Table 6 controls for additional wage bill lags: 2008Q1, 2008Q2, 2008Q3, and 2008Q4. We estimate, using this set of controls, that a $1 injection of spending has a direct wage bill effect of $0.98 (SE = 0.23) and a spillover effect of $0.48 (SE=0.20).

In column (3) we account for the three year moving average of per capita personal income from 2005 to 2008, the share of employment in manufacturing, the log of population, and the change in the unemployment rate between January 2008 and January 2009. Our results change only a small amount.[18] The direct wage bill effect is estimated as $0.75 (SE = 0.22) and the spillover effect is

[18] Twenty-nine observations are dropped because manufacturing employment data were unavailable due to confidentiality concerns. See online QCEW documentation for an explanation of reporting procedures.

Table 6: Two-stage least squares estimates of the direct and spillover effects on the wage bill of Recovery Act spending

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.88***	0.98***	0.75***	0.64***	0.55**
($1 million p.c.)	(0.26)	(0.23)	(0.22)	(0.22)	(0.22)
Adjacent ARRA expenditure	0.46***	0.48**	0.24	0.50***	0.42***
($1 Million p.c.)	(0.15)	(0.20)	(0.17)	(0.17)	(0.15)
Wage bill level (2007Q4)	-0.72***	-0.52	0.12	0.05	0.21
	(0.05)	(0.71)	(0.68)	(0.61)	(0.59)
Income	-	-	-41.27***	-40.52***	-45.89***
(3-yr moving average)†			(9.74)	(8.75)	(8.89)
Log of population†	-	-	0.00	0.01*	0.01**
			(0.00)	(0.00)	(0.01)
Manufacturing share†	-	-	-0.06**	-0.07***	-0.18***
			(0.02)	(0.02)	(0.05)
Change in the Unemployment	-	-	-0.02***	-0.02***	-0.01***
Rate, Jan. 2008 to Jan. 2009			(0.00)	(0.00)	(0.00)
Proportion of Employment in	-	-	-	-	0.13***
Tradable Sector					(0.04)
Log Change in FHFA HPI, 2002Q4	-	-	-	-	-0.14***
to 2005Q4					(0.03)
Log Change in FHFA HPI, 2005Q4	-	-	-	-	-0.12***
to 2009Q4					(0.03)
Added Wage Bill Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Wage Bill Lags	No	No	No	Yes	Yes
N	1630	1630	1601	1601	1601
Kleibergen-Paap F-statistic	47.498	54.783	54.476	69.368	66.091

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
†Coefficients/SEs are rescaled by 100 to ease interpretation.
* $p < .1$, ** $p < .05$, *** $p < .01$

estimated as \$0.24 (SE = 0.17).

Column (4) of Table 6 contains the benchmark controls. The coefficients (ψ^D, ψ^S) equal (\$0.64, \$0.50). Note that the direct and spillover effect estimates are each statistically significant. Also and perhaps not surprisingly, the direct effect of spending is greater than the spillover effect. By summing the two effects, we derive a combined wage bill increase of \$1.14 for every \$1 spent within a subregion.

Next, we show that adding additional control variables has almost no effect on our estimates. In moving from column (4), our benchmark model, to column (5), we add two house price growth variables and the proportion of workers in the tradable goods sector. The coefficients on the direct and spillover effect change very little, even though each new conditioning variable has a statistically significant effect on the wage bill. The stability of our benchmark estimate to adding additional controls provides additional support for our conditional exogeneity identification approach.

The least squares estimates analogous to the estimates in Table 6 are reported in Table 7. By design, some Recovery Act components were intended to be allocated to those regions of the country, or local labor markets, most severely affected by the recession. If funds were ultimately allocated in this manner, than the least squares model should produce causal impact estimates biased in the downward direction. Rather, we find that total effect (direct plus spillover) from the least squares model and the 2SLS model are quantitatively similar. Column (4) of Table 7 reports the least squares specification analogous to our benchmark model, in which we find that each \$1 of Recovery Act spending was associated with a total increase in the wage bill of \$1.16. Compare this to our benchmark result: each \$1 of aid led to an increased regional market wage bill of \$1.14.

This similarity suggests that the allocation of Recovery Act aid was not endogenous to the severity of the downturn at the local labor market level. We note that this finding is consistent with the work of Boone, Dube and Kaplan (2014) who find no relationship between the magnitude of a congressional district's economic downturn and the amount of Recovery Act aid it received.[19]

It should be noted that the wage bill increase resulting from Recovery Act funding does not fully reflect the full value of dollars paid out in the form of contracts, grants, and loans. We cannot rule out the possibility that some dollars went to other regions through cross-regional trade. In Subsection 5.1, we consider the combined regional effect of spending at varying levels of aggregation to provide evidence that our method for identifying regional markets mitigates this confounding form of "leakage."

[19] Although Boone, Dube and Kaplan (2014) show that aid was uncorrelated with the *severity* of the economic downturn, they do find that the level of employment and the poverty rate are both strong predictors of the receipt of aid. This is not problematic for our study. In our baseline specification we include lagged values of employment per capita, which is the more important predictor according to Boone, Dube and Kaplan (2014); furthermore, including the poverty rate as an additional control does not alter our baseline estimates. If employment per capita in 2008 was the primary mechanism by which aid was allocated, then the residual variation in overall aid can be interpreted as, in some sense, unanticipated aid. The similarity between our 2SLS and OLS results reinforces this interpretation.

Table 7: OLS estimates of the effect on the wage bill of Recovery Act spending, aggregate results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	1.01***	1.03***	0.95***	0.90***	0.88***
($1 million p.c.)	(0.15)	(0.21)	(0.22)	(0.21)	(0.20)
Adjacent ARRA expenditure	0.18**	0.20**	0.01	0.16**	0.15**
($1 Million p.c.)	(0.07)	(0.08)	(0.05)	(0.07)	(0.07)
Wage bill level (2007Q4)	-0.75***	-0.52	0.14	0.10	0.26
	(0.05)	(0.73)	(0.68)	(0.61)	(0.59)
Income	-	-	-40.73***	-39.42***	-44.21***
(3-yr moving average)†			(9.84)	(9.02)	(9.21)
Log of population†	-	-	0.01	0.01*	0.01**
			(0.00)	(0.00)	(0.01)
Manufacturing share†	-	-	-0.06**	-0.07***	-0.17***
			(0.02)	(0.02)	(0.05)
Change in the Unemployment	-	-	-0.02***	-0.02***	-0.01***
Rate, Jan. 2008 to Jan. 2009			(0.00)	(0.00)	(0.00)
Proportion of Employment in	-	-	-	-	0.13***
Tradable Sector					(0.04)
Log Change in FHFA HPI, 2002Q4	-	-	-	-	-0.14***
to 2005Q4					(0.03)
Log Change in FHFA HPI, 2005Q4	-	-	-	-	-0.12***
to 2009Q4					(0.02)
Added Wage Bill Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Wage Bill Lags	No	No	No	Yes	Yes
N	1630	1630	1601	1601	1601
R^2	0.476	0.651	0.709	0.725	0.734

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
†Coefficients/SEs are rescaled by 100 to ease interpretation.
* $p < .1$, ** $p < .05$, *** $p < .01$

3.3 Employment response estimates

Increasing employment was, perhaps, the primary objective of the Recovery Act. With this in mind, we re-estimate (3.1), but we replace the wage bill with employment as our outcome variable.

Our specific outcome variable is the accumulated number of job-years, relative to 2008Q4 over the first two years following the act's passage. The formula is:

$$\Delta\text{Job-years}_{j,s} = \frac{1}{4 \times (Pop_{j,s} + Pop_{j,-s})} \sum_{k \in K} (Y_{j,s,k} - Y_{j,s,2008Q4}), \tag{3.2}$$

where $K = \{2009Q1, ..., 2010Q4\}$. One should interpret the coefficient of interest as the number of jobs created and saved (lasting one year each) for every million dollars of Recovery Act money spent.

The sets of conditioning variables used in the job-years regressions are similar, though not identical to those used previously. Instead of using the wage bill trend variables, we use various pre-recession lags of the total number of jobs in each subregion per capita.[20]

In column (4), our benchmark specification, we observe a direct effect of 10.26 (SE = 3.84), which implies that employment increased by roughly 10.26 jobs for every million dollars spent by the federal government. The indirect effect equals 8.50 (SE = 2.81). Both are statistically different from zero at a 1% confidence level. As one might expect, the direct effect on employment is larger than the spillover effect.

From these estimates, we can compute a dollar cost of creating a job by combining the direct and spillover effects.[21]

The implied cost-per-job estimate is $53,305 (=(1e+6)/(10.26+8.50)). To know whether $53,305 is a high or low cost, we compare this number to the typical U.S. worker's earnings. Clearly, this amount depends upon whether a worker is full or part time. Suppose that the Recovery Act created/saved jobs in the same proportion as the numbers of part and full time jobs in the U.S. economy overall. Then it would be appropriate to compare $53,305 to the earnings of a typical worker (averaged across full and part time). This latter number is $40,200.[22] Taken together, we observe that the cost to the government to add a single job for a year was approximately 33% higher than the typical compensation for a single job in the economy overall.

The least squares estimates analogous to those in Table 8 are reported in Table A.1 in the appendix. As with the wage bill results, the employment response estimates from the least squares

[20] As in the wage bill specification, these trend variables are scaled by the population in the regional market as a whole.

[21] More specifically, this is the cost of creating a job-year, which is a job lasting one year.

[22] We compute this number based on the following evidence. According to the 2009 Occupational Employment Statistics, the median hourly wage was $15.95 in 2009. According to the Current Employment Statistics, the average hours worked per week in 2009 was 33.9. The Employer Cost for Employee Compensation reported that wages accounted for 70% of total compensation. Assuming a 52-week work year, this implies a typical annual employment compensation $40,200.

Table 8: Two-stage least squares estimates of the direct and spillover effects on employment of Recovery Act spending

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	18.16***	13.44***	9.94***	10.26***	9.28**
($1 million p.c.)	(5.55)	(3.91)	(3.73)	(3.84)	(4.00)
Adjacent ARRA expenditure	8.56***	5.84**	3.88	8.50***	7.55***
($1 Million p.c.)	(2.51)	(2.78)	(2.68)	(2.81)	(2.57)
Job level (2007Q4)	-0.09***	-0.40***	-0.35***	-0.25**	-0.22*
	(0.01)	(0.13)	(0.13)	(0.12)	(0.12)
Income	-	-	-616.64***	-548.16***	-639.05***
(3-yr moving average)†			(167.72)	(144.38)	(140.13)
Log of population†	-	-	0.06	0.11**	0.16***
			(0.04)	(0.05)	(0.06)
Manufacturing share†	-	-	-0.64	-0.76	-1.80**
			(0.48)	(0.46)	(0.88)
Change in the Unemployment	-	-	-0.20***	-0.22***	-0.21***
Rate, Jan. 2008 to Jan. 2009			(0.03)	(0.04)	(0.04)
Proportion of Employment in	-	-	-	-	1.40**
Tradable Sector					(0.67)
Log Change in FHFA HPI, 2002Q4	-	-	-	-	-2.24***
to 2005Q4					(0.47)
Log Change in FHFA HPI, 2005Q4	-	-	-	-	-1.54***
to 2009Q4					(0.39)
Added Employment Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Employment Lags	No	No	No	Yes	Yes
N	1630	1630	1601	1601	1601
Kleibergen-Paap F-statistic	55.465	56.307	54.205	63.977	62.220

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
†Coefficients/SEs are rescaled by 100 to ease interpretation.
* $p < .1$, ** $p < .05$, *** $p < .01$

model are similar to those from the 2SLS specification. This suggests Recovery Act aid was to a large extent conditionally exogenous to the severity regional downturns. Finally, the first stage results for the job-years specification are provided in Tables A.2 and A.3 in the appendix.

4 Decomposing spillovers: the dynamic and sectoral breakdown of effects

4.1 The dynamic labor market effects of spending

Next we examine how the direct and spillover effects vary with respect to the time since passage of the act. In our previous specifications, we summed the change in the treatment variable relative to the 2008Q4 baseline over the first 8 quarters following the act's passage and observed our outcome variable over the same horizon. To examine the dynamic impact of stimulus spending, we vary the horizon over which we calculate both the outcome and treatment variables.

Figure 2 contains the wage bill responses, the direct effect (dark red, square markers) and the spillover effect (red, diamond markers), for varying horizons over which the wage bill effect is calculated. For example, examining the rightmost side of the figure, the direct effect on the accumulated wage bill change between 2008Q4 and 2012Q4 is approximately 1 and the corresponding spillover effect is roughly 0.6. Both the direct and spillover responses vary little with the horizon and the direct effect is always larger than the spillover effect. We plot 90% confidence intervals around our point estimates, which show that these dynamic effects at longer horizons are imprecisely measured.

Figure 3 contains the corresponding plot for the job-years effect. It shares the same qualitative features as the wage bill figure.

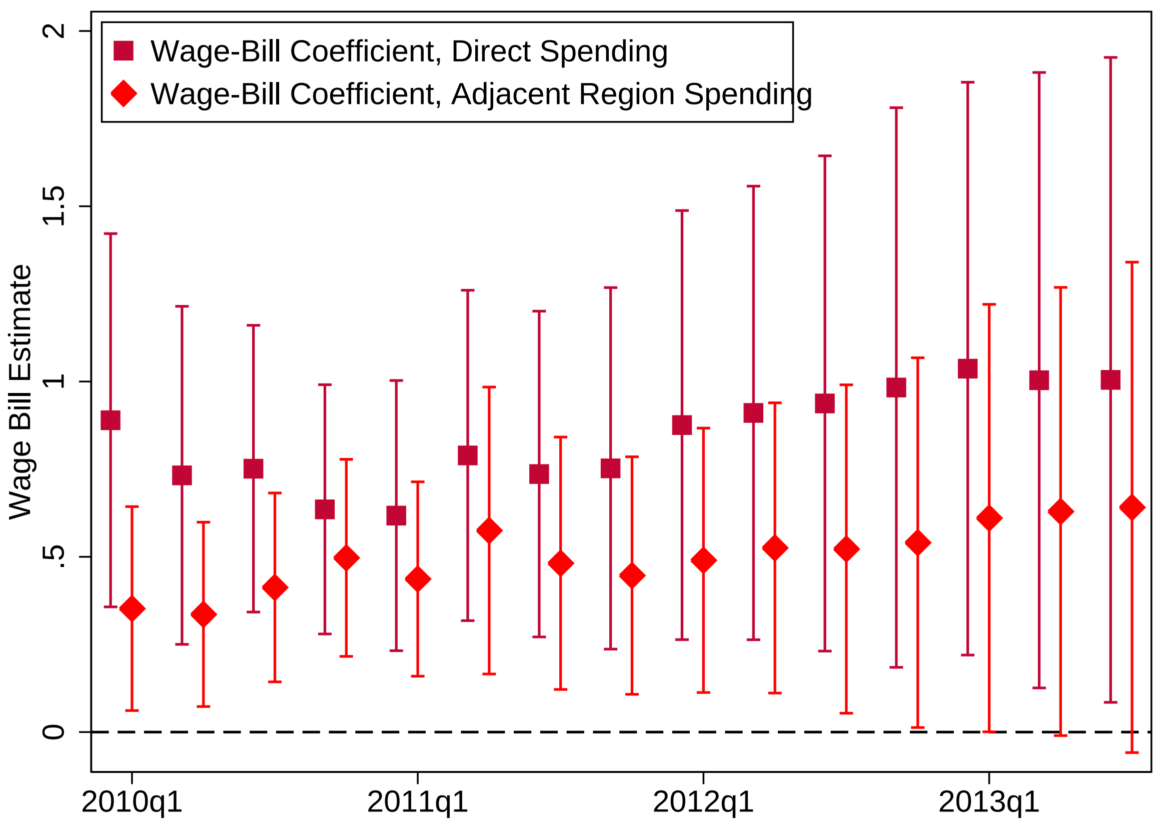
The local benefit to markets receiving stimulus aid appears to be substantial and long-lived. One caveat is in order. To the extent that capital and labor are mobile, our estimates measure not only the immediate benefits accruing to regional markets, but also the endogenous reallocation of economic activity away from other regions of the country towards those receiving such aid. The adjacent spending variable in our regressions controls for reallocation effects to the extent that reallocation is geographically determined; however, it is an open question as to whether other sources of spatial reallocation drive the local multiplier estimates, rather than standard aggregate demand effects.[23]

4.2 Spillovers in the tradable and non-tradable sectors

Our spillover measure is based on commuting patterns. If government purchases in subregion A puts more income into the hands of commuters who work in subregion A but reside in subregion

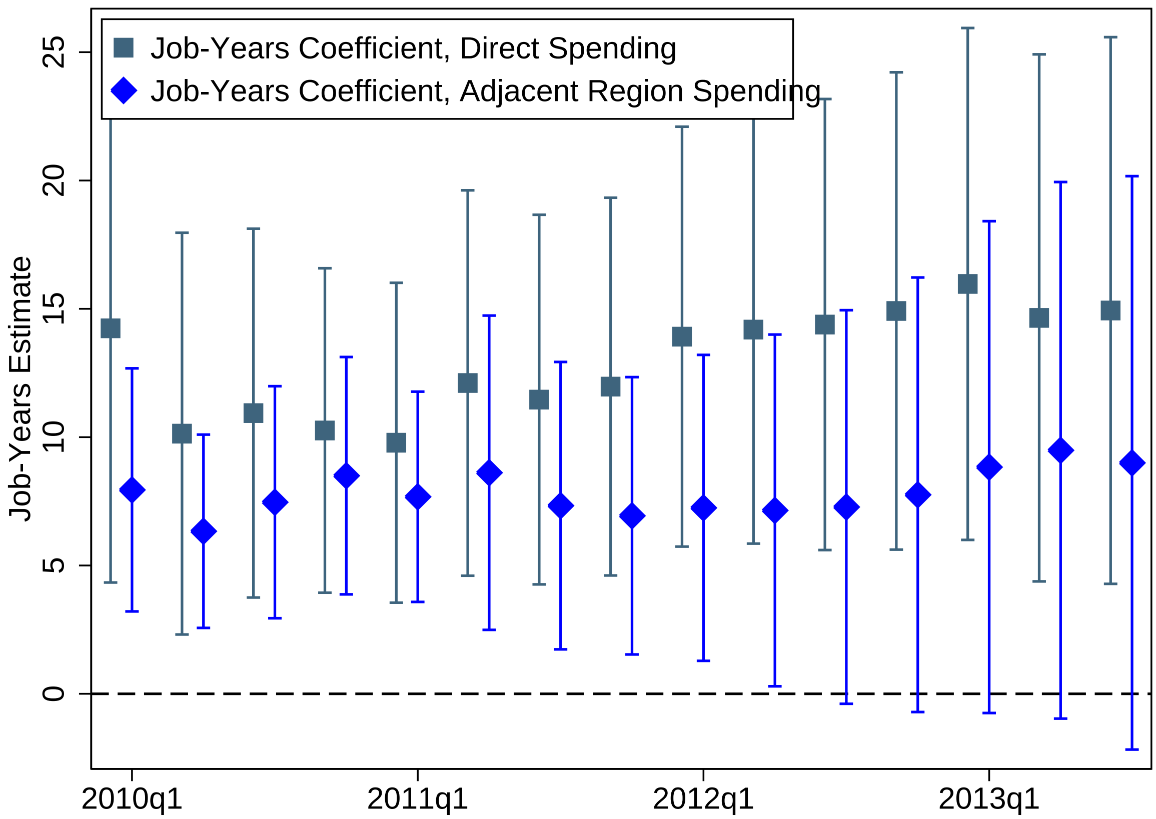
[23] Kline and Moretti (2013) study the long-run effects of the Tennessee Valley Authority and identify long-lived, localized gains in manufacturing—consistent with the presence of agglomeration economies—that were fully offset by losses elsewhere in the country.

Figure 2: Dynamic wage bill effects of spending through 2012Q4, LM_{1293}



Notes: Alaska and regional markets with fewer than 25,000 residents are excluded from the analysis. Equations are estimated by 2SLS with the benchmark set of controls along with Huber-White robust standard errors clustered by state. Error bars correspond to 90% confidence intervals.

Figure 3: Dynamic employment effects of spending through 2012Q4, LM_{1293}



Notes: Alaska and regional markets with fewer than 25,000 residents are excluded from the analysis. Equations are estimated by 2SLS with the benchmark controls along with Huber-White robust standard errors clustered by state. Error bars correspond to 90% confidence intervals.

B, then the spillover effect should manifest itself as greater employment and a higher wage bill in the goods and services that residents of B purchase.

The localized effects of the spillover income are likely to be seen most intensely in the non-traded sector (which is most closely aligned with services in our data set). On the other hand, the direct effect of spending is more likely to be seen in both the goods and services sector, at a minimum, because a substantial part of Recovery Act spending came as goods purchases.

With these channels in mind, we decompose the wage and employment dependent variables into two categories: the goods-producing and service-producing industries.[24] We then re-estimate our benchmark model over three different horizons: 2010Q4, 2011Q4, and 2012Q4. Along these different horizons, we change in tandem both our treatment variable (spending through a given quarter) and our outcome variable (the wage bill and employment responses through a given quarter).

The results appear in Table 9. For a given horizon, the left-side column reports the direct and spillover coefficients for the services sector and the right-side column reports the corresponding coefficients for the goods sector. For every \$1 of aid through 2010Q4, the direct effect on the wage bill in the services sector was an increase of \$0.30 (SE = 0.11) and the spillover effect was an increased wage bill of \$0.27 (SE = 0.14). The direct wage bill response in the goods sector was essentially the same: it increased by \$0.32 (SE = 0.13). However, the spillover wage bill response was less than half of that in the services sector and was statistically indistinguishable from zero: \$0.13 (SE = 0.16). This bears out our thinking that the spillover channel likely operates through the services sector.

At longer horizons, the service-sector spillover result becomes more pronounced. Through 2011Q4, each \$1 of stimulus is associated with an increased wage bill of \$0.28 (SE = 0.16) in the adjacent subregion's services sector. The corresponding response in the goods sector is close to zero: \$0.07 (SE = 0.16). At the 2012Q4 horizon, the spillover wage bill effect in the services sector is \$0.44 (SE = 0.23) and the goods sector effect remains indistinguishable from and close to zero. Table A.7, located in the appendix, contains the employment effects broken down by sector and demonstrates a similar pattern.

4.3 Asymmetrical effects within regional markets

Thus far, we have imposed a symmetry restriction on our estimates: the direct effect of fiscal policy intervention in the large county subregion and in the satellite subregion are constrained to be equal to one another. Similarly, the spillover effects originating from spending in each of the two subregions are likewise constrained to be identical. In what follows, we allow the effects of fiscal policy to vary according to the location within the regional market where the money is spent.

We do this by estimating our benchmark wage bill and job-years models with two sub-samples,

[24] The QCEW provides NAICS sectors codes allowing us to differentiate between Goods-Producing employment and wages and Service-Providing employment and wages at the local level.

Table 9: Two-stage least squares estimates of the wage bill response in the services and goods producing sectors, over various treatment/outcome horizons.

	Thru 2010Q4		Thru 2011Q4		Thru 2012Q4	
	Services Sector	Goods Producing Sector	Services Sector	Goods Producing Sector	Services Sector	Goods Producing Sector
	(1)	(2)	(3)	(4)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.30***	0.32*	0.28	0.46*	0.38	0.58
($1 million p.c.)	(0.11)	(0.17)	(0.18)	(0.25)	(0.27)	(0.38)
Adjacent ARRA expenditure	0.27**	0.13	0.28*	0.07	0.44*	0.02
($1 Million p.c.)	(0.14)	(0.16)	(0.16)	(0.16)	(0.23)	(0.21)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	1601	1601	1601	1601	1601	1601
Kleibergen-Paap F-statistic	69.368	69.368	65.007	65.007	47.255	47.255

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

the large county subregions (columns (1) and (3) of Table 10) and the satellite subregions (columns (2) and (4)). We only look at outcomes within regional markets for which we have non-missing data for both the large county and satellite subregions. The first two columns of Table 10 report the relevant direct and spillover coefficients for the wage bill model and the latter two columns provide the analogous job-years coefficients.

The overall effect of spending in particular subregion can be interpreted as follows. For every $1 spent by the federal government in the large county subregion, its own wage bill increased by $0.84 (SE = 0.30) and the wage bill in the satellite subregion increased by $0.34 (SE = 0.13). The total effect on the regional market is thus $1.18 for every $1 spent in the largest county. Similarly, the overall effect of spending within the satellite subregion can be extrapolated by combining its direct effect of $0.34 (SE = 0.22) with its spillover effect upon the large county subregion of $0.93 (SE = 0.30). For every $1 spent within the satellite subregion the wage bill in the entire regional market increased by $1.27.

Performing similar calculations for the job-year regressions, we observe that for every $1 million spent in the large county subregion there were 12.60 jobs added/saved within the regional market compared with 30.79 jobs added/saved within the entire regional market in response to $1 million spent in the satellite subregion.

Table 10: Two-stage least squares estimates of the wage bill and employment response by location of fiscal policy intervention

	Wage Bill		Job-Years	
	Large County Subregion	Satellite Subregion	Large County Subregion	Satellite Subregion
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.84***	0.34	10.01**	13.04**
($1 million p.c.)	(0.30)	(0.22)	(4.66)	(5.82)
Adjacent ARRA expenditure	0.93***	0.34***	17.75***	2.59
($1 Million p.c.)	(0.30)	(0.13)	(3.28)	(2.78)
All Controls	Yes	Yes	Yes	Yes
N	786	786	786	786
Kleibergen-Paap F-statistic	75.683	80.306	65.396	70.670

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

5 Robustness: Choice of LM_J, placebo diagnostics, model specification, outlier leverage, instrument validity

5.1 Comparison across choices of LM_J

How does our choice of LM_{1293} relate to estimates derived from other possible levels of aggregation? The level of aggregation is likely to be important because as we consider successively more aggregated data, we will be combining more and more counties. In combining counties, we will be subsuming the spillover effects between those counties. The subsumed spillover effects should then manifest themselves as part of the estimated direct effect.

As such, varying the degree of aggregation should be sufficient for us to see a spillover effect. To this end, we restrict ψ^S and ν^S to each be equal to zero and redefine a single observation to be the combination of the large county subregion and outlying subregion pair.

Then, we estimate the model for different degrees of aggregation. The results are presented in Figure 4a. On the horizontal axis, we list the number of local labor markets at a particular level of aggregation. For example, on the rightmost tick on the x-axis we have 3,144, which corresponds to county level data (i.e. no aggregation). As one moves from right to left, the degree of aggregation becomes successively stronger. The leftmost tick has the U.S. combined into 372 distinct local labor markets. This is approximately 8.5 counties per region. Values on the vertical axis indicate the wage bill response to $1 of Recovery Act spending within a region. The solid line indicates the point estimate of the effect and the dashed lines envelope the 90% confidence intervals.

At the most disaggregate level, LM_{3144}, the point estimate is 0.39 and is precisely estimated. As the level of aggregation increases from LM_{2500} to LM_{1293}, the wage bill effect increases. At LM_{1293}, we can reject the county-level wage bill response of 0.39 at the 99% confidence level.[25]

We interpret this part of the figure to mean that, as we aggregate from LM_{3144} to LM_{1293}, additional positive spillover effects are being subsumed because we are including additional counties into each observation.

To the left of LM_{1293} there is suggestive evidence of additional spillovers between local labor market that are not accounted for in the LM_{1293} classification. The point estimate on the wage bill response is $1.12 and increases somewhat smoothly to $1.87 by the time we aggregate to LM_{372}. However, we lack the statistical precision to reject the $1.12 estimate in our benchmark specification. Hence, the LM_{1293} environment provides a conservative framework (i.e. biasing ourselves against finding large spillovers) in which to understand spatial spillovers within local labor markets.

The numbers corresponding to the information in Figure 4a appear in Table 11, for select LM_J.

Our regional market approach, as explained above, incorporates potential cross-county spillover effects into the analysis. A cross-sectional analysis that ignores such spillovers may yield biased estimates. The county-level analysis yields a point estimate of 0.39 (SE = 0.19). If taken at face value, this implies that Recovery Act funding was less than half as effective at increasing the wage bill as our benchmark results imply. Given that our regional market estimate is quantitatively large, the difference is stark and the implied bias from ignoring such spillovers is thus quite substantial.

Figure 5a shows the corresponding results of the job-years regressions.[26] As with the wage bill figure, the coefficient increases with the degree of aggregation up until LM_{1293}. At this point and at higher levels of aggregation, the point estimate stabilizes at approximately 20. Beyond LM_{1293}, the spillovers we measure appear to be fully contained within the identified regional markets and as such there are little additional spillovers between them for us to identify. That is, no new information is gained by agglomerating regions further.

Recall that, in Section 3.3, the combined direct and spillover employment effect in our benchmark specification was 18.8. Thus, the results from the two approaches are consistent with each other.

[25] In unreported regressions we estimate the county-level model except that we include all counties regardless of population size and find a statistically insignificant, near-zero effect of ARRA spending upon the wage bill and employment.

[26] All of these estimates include the full set of conditioning variables. Tabulations corresponding to Figure 5a of select LM_J are provided in the appendix.

Table 11: Two-stage least squares estimates of the wage bill response along different degrees of aggregation

	County Level	LM_{1293}	LM_{601}	LM_{372}
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.39**	1.12***	1.63***	1.87***
($1 million p.c.)	(0.19)	(0.25)	(0.32)	(0.48)
All Controls	Yes	Yes	Yes	Yes
N	1587	918	510	330
Kleibergen-Paap F-statistic	26.127	132.380	37.116	22.797

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

5.2 Placebo Diagnostics

We perform two placebo diagnostics to assess whether the commuting spillovers previously estimated might be spuriously estimated.

Shuffled Local Labor Markets

The first diagnostic is performed using randomly determined local labor markets. First, we group counties into quintiles according to their employed resident labor force. Then, within each quintile, we randomly reassign the geographic location of each county to that of a different county from the same quintile. Apart from the change in location, each county keeps all of its own variables used in the analysis (e.g. employment, wage bill, Recovery Act spending). Then as before, for each level of aggregation, we calculate the regional variables. Relative to the "true" non-shuffled geographies, our procedure maintains the total number of regional markets, the number of counties comprising each region and, roughly, the labor force distribution within each region. However, because we have randomized locations, we have broken the commuting linkages within regional markets in the original data.

Next, we estimate the regional 2SLS model and plot the estimate for the shuffled data at various levels of aggregation. If the spillovers we identify are strongly tied to commuting linkages between counties (such as we see in Figure 4a), then we would expect our wage bill estimate to be unrelated with the degree of randomized aggregation. As seen in Figure 4b, this is precisely what happens. Using the shuffled data, our wage bill response estimate is flat with respect to the degree of aggregation.

Figure 5b presents the analogous estimates for the employment responses. The conclusion for this variable is the same.

Shuffled Spillover Treatment

In our second placebo diagnostic exercise, we explore more directly how the spillover estimates change (if at all) when the spillover treatment is randomly determined, holding all other aspects of the benchmark model constant.

We randomize the spillover treatment for the large county subregion from LM_{1293} as follows. First, we group these large county subregions into quintiles according to population. Within each of these quintiles, we randomly reassign to each large county subregion the total value of stimulus and composite instrument spending through 2010Q4 actually received by another large county subregion.[27]

For each large county subregion, there is a corresponding satellite subregion. For this satellite subregion, we construct the spillover ARRA treatment in an identical fashion as in the benchmark model, except that the numerator of the spillover variables ($ARRA_{j,1}$ and $CompARRA_{j,1}$) are taken to be the randomly determined values of spending assigned to the large county subregion. We construct a similar randomized spillover treatment variable for the large county subregions by following the same aforementioned steps with the satellite subregion instead.

Using the randomly determined spillover treatment, we estimate what would otherwise be the benchmark model. Thus, we sever the connection between outcomes in the labor market (wage bill and employment) and Recovery Act spending in the adjacent subregion.

Table 12 compares the results of this placebo exercise with our benchmark results. Columns (1) and (3) reproduce our benchmark estimates for the wage bill and employment models, respectively. Columns (2) and (4) contain the randomized spillover placebo results. Observe first the results from the wage bill random placebo model in column (2). Whereas the direct effect of every $1 of aid is nearly identical to the benchmark model (an increased wage bill of $0.66), the effect of randomly determined aid to an adjacent subregion is close to and statistically indistinguishable from zero. This is precisely as we would expect since the adjacent subregion ARRA aid no longer represents the actual value that subregion received. The outcome is the same when we use the employment variable.

5.3 Alternative Specifications

Table 13 provides estimates of the wage bill subregion model for alternative specifications. Column (1) contains the benchmark specification. Column (2) contains the population-weighted, as opposed to the benchmark case which is unweighted. For the weighted estimate, the spillover effect coefficient

[27] It is possible that a particular county is randomly paired with its actual treatment; however, this occurs in less than 1% of all cases.

Table 12: Two-stage least squares estimates of the employment and wage bill response with randomized placebo spillover treatment

	Wage Bill		Job-Years	
	Benchmark	Random Placebo	Benchmark	Random Placebo
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.64***	0.66***	10.26***	10.63***
($1 million p.c.)	(0.22)	(0.22)	(3.84)	(3.88)
Adjacent ARRA expenditure	0.50***	-	8.50***	-
($1 Million p.c.)	(0.17)		(2.81)	
Random ARRA expenditure	-	0.02	-	0.15
($1 million p.c.)		(0.03)		(0.45)
All Controls	Yes	Yes	Yes	Yes
N	1601	1601	1601	1601
Kleibergen-Paap F-statistic	69.368	62.591	63.977	56.713

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state. The randomized placebo results are identical to the benchmark specification except that the numerator in the spillover treatment variable is randomly determined to be identical to that of another subregion of the same type (large county or satellite subregion) and same quintile of population.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 13: Two-stage least squares estimates of the wage bill response for various specifications

	Benchmark	Weighted	LM_{372}	LM_{372} $Weighted$
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.64***	0.74*	0.83**	0.85*
($1 million p.c.)	(0.22)	(0.39)	(0.35)	(0.48)
Adjacent ARRA expenditure	0.50***	0.01	0.95***	0.21
($1 Million p.c.)	(0.17)	(0.33)	(0.31)	(0.45)
All Controls	Yes	Yes	Yes	Yes
N	1601	1601	655	655
Kleibergen-Paap F-statistic	69.368	37.571	21.571	26.985

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

is approximately zero while the direct effect coefficient increases by a small amount relative to the benchmark coefficient.

Why might the spillover effect disappear in the weighted specification? One possibility is that subregions from large regional markets are more likely to be self-sufficient than small ones. If a large subregion's economy is sufficiently diversified in terms of the goods and services it creates and the skills of its workforce, then the subregion's interconnectness with nearby areas may engender a weak transmission mechanism for spillovers. Weighting by population, in turn, would make it difficult to identify the spillovers if they are more prevalent in low population local labor markets.

Column (3) contains the unweighted estimates where the number of local labor markets equals 372 instead of our benchmark 1293. As with the benchmark specification, there is a statistically significant direct and spillover effect. This demonstrates that the finding of spillover effects across subregions was not contingent on the particular level of disaggregation we initially selected. Note that both effects are somewhat larger than their counterparts from the benchmark specification and the standard errors increase substantially relative to benchmark case. Column (4) contains the weighted estimates at the LM_{372} aggregation level. The direct effect coefficient is statistically significant, but the spillover effect coefficient is not. However suggestive, a difficulty with this specification is that the standard errors are much larger than those from the benchmark specification.

Table A.5 provides the corresponding estimates to Table 13 except where the outcome variable is job years. The takeaways are similar to those in the previous table. First, using the 372 LLM aggregation level results in little change in the job-year estimates relative to the benchmark case. Second, weighting by population, both at the 1293 and 372 level of disaggregation, results in a dramatic drop in the spillover coefficient.

5.4 Outlier leverage

As an additional robustness check, we explore the importance of outliers. First, we compute predicted values for direct and spillover spending from the first stage exactly as in the benchmark specification.

Then, we regress direct ARRA spending upon all other conditioning variables (including the predicted spillover ARRA spending), and define the resulting residuals to be the conditional ARRA variation. Similarly, we regress our outcome variable upon all conditioning variables (including predicted spillover ARRA spending) excluding direct ARRA spending, producing residuals which we denote as conditional outcome variation. We do this for each outcome variable.

At this point, regressing conditional ARRA variation upon conditional outcome variation would produce an estimate of ψ^D that is identical to what is reported in Table 6 as the direct effect of ARRA spending. Because we are interested in removing any possible influence of outliers, at this juncture we winsorize each of the two series at the 1% and 99% level.[28] This entire procedure, *mutatis mutandi*, is performed with the predicted spillover ARRA series as well.

Figure 6 contains the four derived winsorized partial regression plots. The top row produces binned scatter plots of the direct (left) and spillover (right) effects of spending upon the wage bill. The bottom row produces the analogous estimates for employment. There is a discernible, albeit noisy, linear relationship between stimulus and employment and wages. Furthermore, we note that when we limit the leverage of outlier values, both the direct and spillover estimates increase by a moderate amount relative to our benchmark results. This exercise produces the following coefficients (benchmark estimates in parentheses): for every $1 million of aid there was a direct effect upon the wage bill of $0.80 ($0.64), a spillover effect upon the wage bill of $0.64 ($0.50), a direct effect upon employment of 13.40 job-years (10.26 job-years), and a spillover effect upon employment of 11.41 job-years (8.5 job-years).

These moderately increased coefficients are well within a single standard deviation away from the benchmark results.

5.5 Assessing instrument validity

An alternative instrument

Our basic endogeneity concern is that governments targeted Recovery Act funds to areas worst affected by the recession. Our instrument consists of spending by agencies which were not directed to allocate funds according to this criteria. One might be concerned that, for various reasons, these agencies still chose to use this criteria for allocating funds.

[28] Any value contained in the vector of conditional ARRA variation that is above the 99th percentile of the distribution is replaced with the value of conditional ARRA variation at the 99th percentile; likewise for values below the 1st percentile. This is done for both conditional ARRA variation and conditional outcome variation.

For this reason, we develop an alternative instrument that uses a different justification. First, we observe that the Recovery Act data tracks payments made by primary recipients to vendors and sub-recipients to sub-vendors. Many of these payments were made to vendors and sub-vendors outside of the local labor market in which the purchasing recipients and sub-recipients resided.

We use the total value of these payments as an alternative instrument, which we call the vendor instrument. For example, suppose a primary recipient located in region X purchased \$50,000 in goods from a vendor located in region Y. This transaction would count as \$50,000 towards spending in region Y in the tabulation of the vendor instrument.

Even if recipients were targeted because of the severity of their local downturn, we contend that there is no reason that a recipient would purchase from a vendor based upon the severity of the recession in the vendor's region. We then scale the total value of all vendor payments made to a subregion to be millions of dollars per person.

Column (1) of Table 14 gives the estimate of the wage-bill direct and spillover estimates using the vendor instrument. Besides the change in instrument, the model is identical to the benchmark specification. The direct effect coefficient equals \$1.20 (SE = 0.45) and the spillover effect coefficient equals \$0.27 (SE = 0.15).[29]

Recall that the corresponding benchmark estimates (using the composite instrument) were \$0.64 and \$0.50, respectively. Quantitatively, the vendor instrument produces a larger point estimate on the direct effect of fiscal aid and a smaller spillover effect than our benchmark estimates. However, these differences lack statistical precision. Observe that the lower bound on the 90% confidence interval in column (1) for the direct effect is \$0.46. That is, accounting for the imprecision of the estimates, the vendor instrument produces an estimate of the direct effect that is statistically indistinguishable from \$0.64—our benchmark estimate. A similar observation holds for the spillover estimate when using the vendor instrument.

Column (2) gives the estimate of the job-years direct and spillover estimates using the vendor instrument. The direct effect coefficient equals 14.9 and the spillover effect coefficient equals 5.05. The corresponding benchmark estimates were 10.25 and 8.49, respectively.

Columns (3) and (4) of Table 14 report the estimates when we include both the composite instruments and the vendor instruments in the estimation. The resulting estimates are similar to those in column (1) and (2). This suggests that our alternative instrument does not a tell story that conflicts with the message delivered by our baseline, composite instrument.

The vendor instrument, which is based on an entirely different justification from the composite instrument, delivers positive direct and spillover effects of Recovery Act spending which are statistically different from zero and economically important. There are some quantitative differences across the results from the two respective sets of instruments; however, these differences are imprecisely measured. We view the qualitative similarities as additional support for the validity of our

[29] The partial F-statistic from a first-stage regression for the vendor instrument is reported at the bottom of column (1). It equals 116, which indicates that the instrument is highly correlated with the endogenous variable.

Table 14: Two-stage least squares estimates of the employment/wage-bill response with vendor payments as the instrument.

	Wage Bill Instrument: Vendor Payments	Job-Years Instrument: Vendor Payments	Wage Bill Instruments: Vendor Payments and Composite	Job-Years Instruments: Vendor Payments and Composite
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	1.20***	14.93***	1.05***	13.75***
($1 million p.c.)	(0.45)	(4.14)	(0.40)	(3.90)
Adjacent ARRA expenditure	0.27*	5.05**	0.34***	6.31***
($1 Million p.c.)	(0.15)	(2.54)	(0.11)	(1.85)
Income	-37.61***	-527.76***	-38.34***	-532.26***
(3-yr moving average)†	(9.08)	(143.00)	(9.01)	(143.09)
Log of population†	0.01*	0.11**	0.01*	0.11**
	(0.00)	(0.05)	(0.00)	(0.05)
Manufacturing share†	-0.06***	-0.69	-0.07***	-0.71
	(0.02)	(0.45)	(0.02)	(0.46)
Change in the Unemployment	-0.02***	-0.22***	-0.02***	-0.22***
Rate, Jan. 2008 to Jan. 2009	(0.00)	(0.04)	(0.00)	(0.04)
Census Region Dummies	Yes	Yes	Yes	Yes
Employment Lags	No	Yes	No	Yes
Wage Bill Lags	Yes	No	Yes	No
Spillover Employment/Wage Lags	Yes	Yes	Yes	Yes
N	1601	1601	1601	1601
Kleibergen-Paap F-statistic	116.418	105.482	27.279	25.128
Hansen J Statistic			0.406	0.466

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

identification approach.

Instrument construction permutations

The instrument's validity is predicated upon each of the federal agency spending totals used to construct our instrument being allocated in a conditionally exogenous manner with respect to local economic conditions. Suppose this assumption is invalid for a subset of the agencies. Using only these agencies to build a new instrument and then re-estimating the model using this instrument should produce biased estimates.

Along the same lines, suppose the conjectured endogenous components are small in dollar amount or degree of endogeneity. Then, if we exclude them from our instrument construction and re-estimate the model, we will find estimates close to our benchmark results.

This leads naturally to the following exercise. There are nine components (i.e., federal departments/agencies) used to construct the instrument, implying 511 possible combinations of components. For each of these 511 possible instrument constructions, we re-estimate the wage bill and job-years regressions. Because there is the additional concern that a particular permutation will produce a weak instrument, we exclude those estimates with Kleibergen-Paap F-statistics below the Stock and Yogo (2005) 10% critical value of 7.03. There are 21 such permutations. From the remaining 490 estimates, we trim the top and bottom percentile and construct kernel density plots.

Figures 7a and 7b provide the results of this instrument permutation exercise. Figure 7a contains the distribution of estimates for the wage bill regression. The dotted vertical lines indicate the coefficients for the benchmark model, i.e., using all nine federal agencies in constructing the instrument. Its value is \$0.64. The blue kernel density plot is for the direct effect coefficient. The $10th$ and $90th$ percentiles of the distribution are \$0.42 and \$0.85. Thus, for the vast majority of potential sets of agencies (from amongst those found by our narrative analysis) that we might have included in the instrument, all of the resulting point estimates would tell the same basic story: The local, direct effect of one dollar of Recovery Act spending is a positive but less than one-for-one increase in the wage bill.

The corresponding distribution from the exercise for the spillover effect (the red line) in 7a is similarly tight. The $10th$ and $90th$ percentile of the distribution based on the instrument permutations are \$0.27 and \$0.70. The benchmark spillover wage bill effect is \$0.50. Once again, most of the potential instrument constructions give similar, although not identical, accounts of the causal impact of Recovery Act spending in one area on neighboring areas.

Figure 7b provides the results of this exercise for the employment model (benchmark coefficient in parentheses): the median direct effect coefficient is 10.09 job-years (10.26 job-years) and the median spillover effect coefficient is 8.46 job-years (8.50 job-years). Thus, our results do not hinge upon an instrument constructed from a conveniently chosen combination of Recovery Act programs.

Pre-act trends and the composite instrument

One can also consider the extent to which the composite instrument is conditionally uncorrelated with pre-Act economic trends. If a strong correlation existed, this might be suggestive that the instrument is endogenous with respect to the error term. Note however that this evidence would be neither necessary nor sufficient to establish endogeneity because what matters in actuality is the correlation between the instrument and the unobservable current shocks.

Nonetheless, we investigate this question by first measuring the pre-Act employment trend as the one-year percentage growth in the sub-regional employment per capita ending in 2008Q4.[30] Then, we estimate a naïve univariate regression of the composite instrument on this trend variable. We call this naïve because in our actual specifications we include several conditioning variables, in part, to soak up any potential endogeneity of the instrument.

The coefficient on the employment trend variable is $0.77 (see column (1) of Table 15). This indicates that 1% faster pre-Act employment growth in a subregion is associated with $0.77 of additional composite instrument aid. This estimate is statistically indistinguishable from zero and economically trivial. Specifically, $0.77 represents only 1.3% of the standard deviation of instrument spending and 2.6% of what the median sub-region received. Along those same lines, very little of the variation in the instrument is explained by the pre-Act trend: The regression's R-squared is less than 1%. Thus, there is no economically significant relationship between component aid and the one-year trend growth in employment.

Moreover, the above regression is not reflective of the fact that our actual regression conditions on additional variables. If we run the same regression as above except we add the employment per capita controls that are included in our benchmark specification, the coefficient on the employment growth trend falls in magnitude and remains statistically indistinguishable from zero (column (2) of Table 15).

6 Conclusion

This paper explores the importance of cross-regional spillovers in assessing the impact of countercyclical government spending. Stimulus spending from the Recovery Act in one county increased employment and wage payments in places two to three counties away, so long as the areas were sufficiently connected as measured by commuting patterns.

The presence of spillovers has important policy implications. For example, we find that when Recovery Act spending took place in a large county, nearly one-half of the resulting increase in employment occurs in surrounding counties. Failing to take into account positive spillovers could lead policymakers to understate the total social benefit of a stimulus program.

[30] The Recovery Act was passed in February of 2009; hence, we do not look at the year-over-year change through 2009Q1 in this analysis. Results are essentially the same when we do.

Table 15: OLS estimates of the association between composite instrument Recovery Act spending and the pre-Act employment trend

	Naïve Coef./SE	Pre-Act Employment Levels Coef./SE
Percent change, employment per capita 2007Q4 to 2008Q4	0.77 (0.63)	0.27 (0.46)
Employment per capita, 2008Q4	-	-44.66 (335.81)
Employment per capita, 2008Q3	-	698.20 (521.04)
Employment per capita, 2008Q2	-	-148.09 (410.22)
Employment per capita, 2008Q1	-	-432.83 (394.07)
Employment per capita, 2007Q4	-	90.36 (444.81)
Constant	45.99*** (3.78)	11.96*** (3.54)
Number of Obs.	1601	1601
R-Squared	0.003	0.106

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

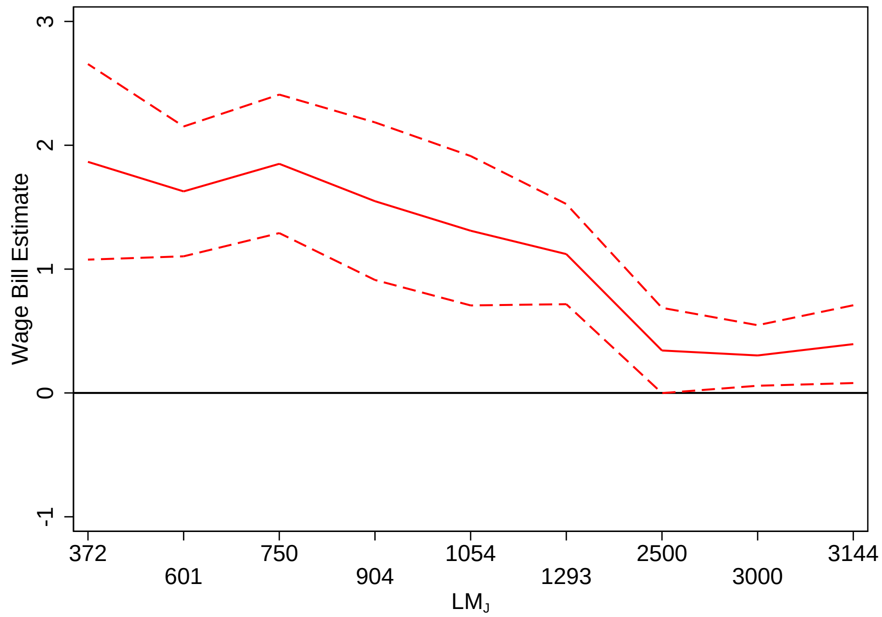
Second, the spillover estimates provide evidence that is consistent with the "Keynesian multiplier" channel. Government spending in one county increases wage payments and employment in other counties. This could be explained by second-round effects of spending that are critical to the traditional Keynesian explanation for how fiscal policy is magnified.

Finally, more work remains to be done since there are other potential sources of spillovers besides the one studied here. For example, individuals consume goods delivered from far outside their areas of residence. These linkages are not captured solely by examining commuting data. Thus, trade in goods may be another source of spillovers.

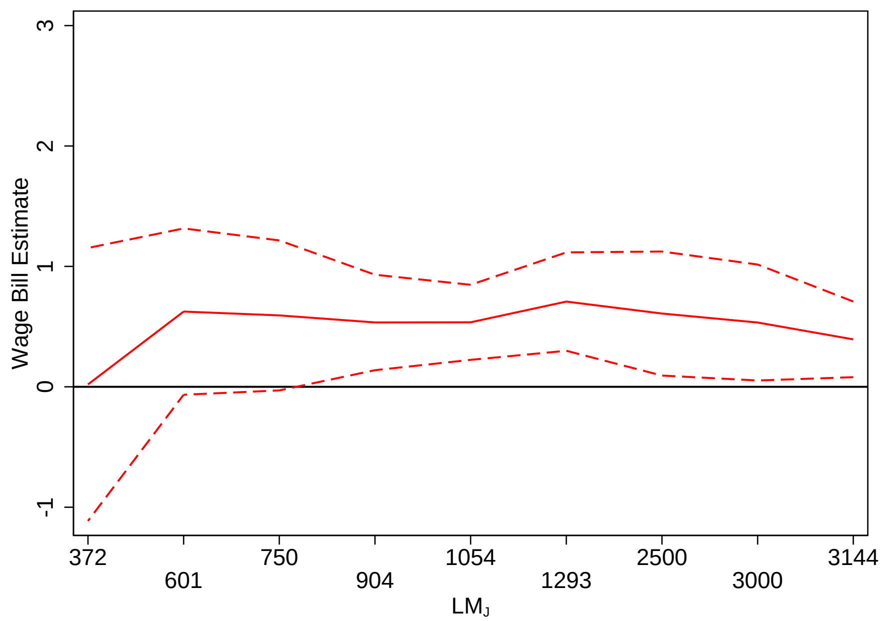
Another potential spillover could exist because the location of government spending may not be the same location at which the taxes to cover the spending will be collected. A Tax Foundation (2005) analysis shows that there is a great deal of heterogeneity across states in the federal spending received per dollar of federal tax dollars paid. As a stylized example, suppose $1 million is spent in Mississippi in one year but the offsetting taxes will eventually be paid by residents in New Jersey. The spending in Mississippi may have a negative spillover on New Jersey if, for example, individuals in the latter state reduce capital accumulation in anticipation of future distortionary taxes.

Figure 4: Estimates of the wage bill effect of government spending for different degrees of aggregation: actual and randomly assigned regional model specification

(a) Wage bill aggregation effects by LM_J



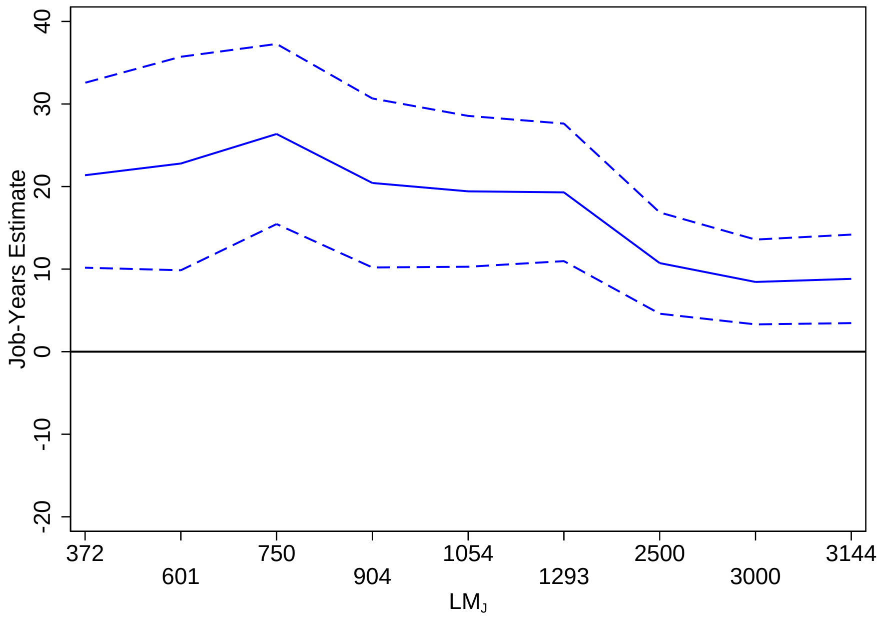
(b) Random assignment wage bill aggregation effects by LM_J



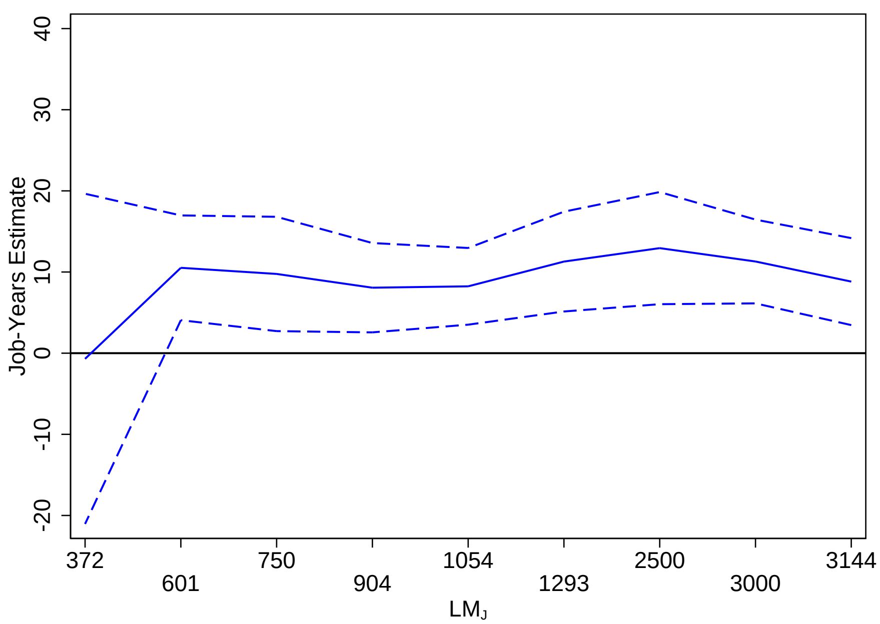
Notes: Alaska and Regional markets with fewer than 25,000 residents are excluded from the analysis. Equations are estimated by 2SLS with the full set of controls along with Huber-White robust standard errors (SEs) clustered by state.

Figure 5: Estimates of employment effect of government spending for different degrees of aggregation: actual and randomly assigned regional model specification

(a) Job-years aggregation effects by LM_J

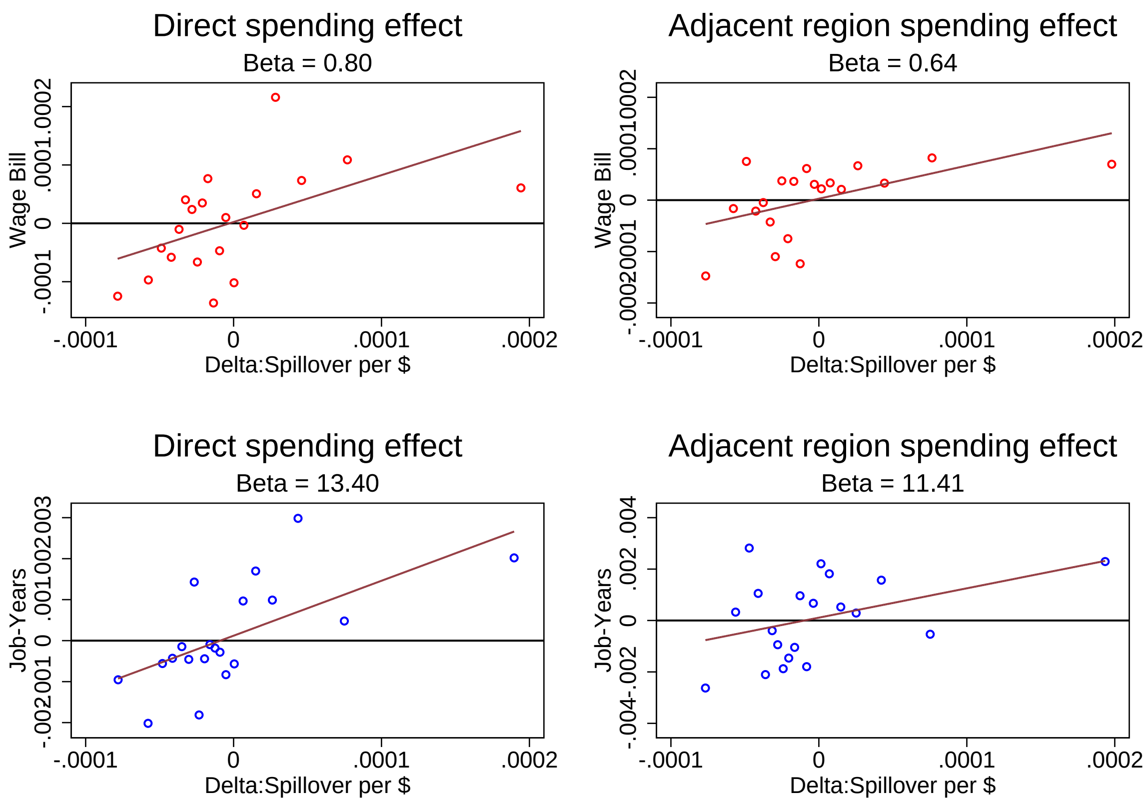


(b) Random assignment job-years aggregation effects by LM_J



Notes: Alaska and regional markets with fewer than 25,000 residents are excluded from the analysis. Equations are estimated by 2SLS with the full set of controls along with Huber-White robust standard errors (SEs) clustered by state.

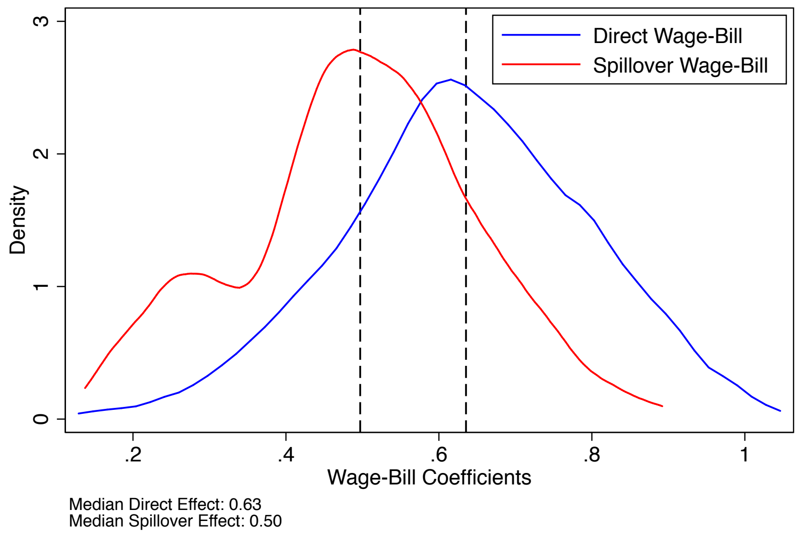
Figure 6: Partial regression plots of wage bill and employment response for direct & spillover estimates from two-stage least squares results, LM_{1293}, with winsorized residuals.



Notes: The estimates above exclude Alaska and regional markets with fewer than 25,000 residents. Equations are estimated by 2SLS with the full set of controls. To make the plots more legible, we place x-axis variables into 20 equal sized bins. We plot, within each bin, the mean value of the x-axis variable against the mean value of the y-axis variable. We also winsorize the residuals at the 1% and 99% level to emphasize the conditionally linear relationship between ARRA funding and the direct and spillover effects.

Figure 7: Kernel density plots of second-stage coefficients for wage bill and job-years treatment effects for 490 possible strong instrument combinations

(a) Wage-Bill Instrument Permutations



(b) Job-years Instrument Permutations

Notes: There are 9 components of our instrument, implying the existence of 511 possible combinations of the various components. We exclude those results with Kleibergen-Paap F-statistics below the 10% critical value of 7.03. Plotted are kernel density plots of the second-stage results. We have trimmed the top and bottom 1% of coefficients.

References

Army Corps of Engineers, U.S. (2010a), "Civil Works Summary Agency Recovery Act Plan," March 31.

Army Corps of Engineers, U.S. (2010b), "Civil Works Program-Specific Agency Recovery Act Plan," March 31.

Autor, D. D. Dorn and G. Hanson (2013), "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121-68.

Beetsma, R. M. Giuliodori (2011), "The Effects of Government Purchases Shocks: Review and Estimates for the EU," *The Economic Journal*, 121(550), F4-F32.

Boone, C. A. Dube and E. Kaplan (2014), "The Political Economy of Discretionary Spending: Evidence from the American Recovery and Reinvestment Act "*Brookings Papers on Economic Activity*, Spring: 375-441.

Bureau of Economic Analysis (2013), "Reconciliation of ARRA Outlays and NIPA Federal Government Statistics."

Carlino, G. and R. Inman (2013), "Macro Fiscal Policy in Economic Unions: States as Agents,"NBER Working Paper 19559.

Chetty, R. N. Hendren, P. Kline, and E. Saez (2014), "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States,"NBER Working Paper 19843.

Chodorow-Reich, G., L. Feiveson, Z. Liscow and W. Woolston (2012), "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, 4(3),118-45.

Clemens, J. and S. Miran (2012), "Fiscal Policy Multipliers on Subnational Government Spending," *American Economic Journal: Economic Policy*, 4(2), 46-68.

Cogan, J. and J. Taylor (2012), "What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package," in: L. and J. Taylor and I. Wright (ed.), *Government Policies and the Delayed Economic Recovery, chapter 5 Hoover Institution*, Stanford University.

Conley, T. and B. Dupor (2013), "The American Recovery and Reinvestment Act: Solely a Government Jobs Program?" *Journal of Monetary Economics*, 60(5), 535-549.

Drautzburg, T. and H. Uhlig (2013), "Fiscal Stimulus and Distortionary Taxation." Federal Reserve Bank of Philadelphia Working Paper 13-46.

Dupor, B. and M. S. Mehkari (2014), "Schools and Stimulus,"Federal Reserve Bank of St. Louis, working paper.

Education, U.S. Dept. of (2009), "Guidance: Funds for Part B of the Individuals with Disabilities Act Made Available under the American Recovery and Reinvestment Act of 2009," July 1, revised.

Energy, U.S. Dept. of (2009), "American Recovery and Reinvestment Act Program Plan for the Office of Energy Efficiency and Renewable Energy," May 15.

Energy, U.S. Dept. of (2010), "American Recovery and Reinvestment Act Program Plan for the Office of Energy Efficiency and Renewable Energy," June 15 (updated).

Environmental Protenction Agency (2009), "Award of Capitalization Grants with Funds Appropriated by P.L. 111-5, the American Recovery and Reinvestment Act," Office of Water, March 2.

Environmental Protenction Agency (2011), "Implementation of the American Recovery and Reinvestment Act of 2009: Clean Water and Drinking Water State Revolving Fund Programs," Office of Water, May.

Federal Register (2009), Federal Register 74 (42), Government Printing Office, 9656-9671.

Galí, J. and T. Montacelli (2005), "Optimal Monetary Policy and Exchange Rate Volatility in a Small Open Economy," *Review of Economic Studies*, 72(252), 707-34.

General Services Administation (2009a), "American Recovery and Reinvestment Act Agency-Wide Recovery Plan."

General Services Administation (2009b), "GSA Motor Vehicle Replacement Plan: Energy Efficient Federal Motor Vehicle Fleet Procurement."

General Services Administation (2012), "Revised American Recovery and Reinvestment Plan #10," U.S. GSA Public Building Services, November 30.

Greenwood, J., Z. Hercowitz and G. Huffman, "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402-17.

Justice, U.S. Dept. of (2009a), "Office of Justice Programs Recovery Act Grants," Office of Justice Programs.

Justice, U.S. Dept. of (2009b), "Recovery Act: Correctional Facilities on Tribal Lands Program Competitive Grant Announcement," Bureau of Justice Assistance, May 4.

Kline, P. and Enrico Moretti (2013), "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." NBER Working Paper 19293.

New America Foundation (2014), "Individuals with Disabilities Education Act-Funding Distribution," Federal Education Budget Project, April.

Shoag, D. (2012), "Using State Pension Shocks to Estimate Fiscal Multipliers since the Great Recession," *American Economic Review*, 103(3), 121-24.

Stock, J. and M. Yogo (2005), "Testing for Weak Instruments in Linear IV Regression," ch. 5 in Donald W.K. Andrews (ed.), *Identification and Inference for Econometric Models*, New York: Cambridge University Press, pp. 80-108.

Suarez Serráto, J. and P. Wingender (2014), "Estimating Local Fiscal Multipliers," Duke University, working paper.

Tax Foundation (2005), "Federal Spending Received Per Dollar of Taxes Paid by State, 2005."

Tolbert, C. and M. Sizer (1996), "U.S. Commuting Zones and Labor Market Areas," Economic Research Service, Rural Economy Division, U.S. Dept. of Agriculture, September.

Wilson, D. (2012), "Fiscal Spending Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, 4(3), 251-82.

A Appendix

A.1 Additional Instrument Construction Information

In the main text, we justify our use of Environmental Protection Agency and Department of Justice Recovery Act spending as components of our instrumental variable. In this section, we provide our reasons for each of the other Recovery Act components being included in the instrument.

General Services Administration (GSA). The Recovery Act provided the GSA with $5.857 billion. Approximately $5.5 billion was appropriated to the Federal Building Fund, to be used to construct and restore federal buildings. Another $300 million was appropriated for the procurement of energy-efficient vehicles in the federal fleet. We use funding from these two components as summands in our instrument amount. General Services Administration (2009a) describes two key goals of its projects: (i) spending money quickly to stimulate the economy and create jobs, (ii) improve the environmental performance of federal assets.

General Services Administration (2009a) states construction projects will take place in all 50 states, the District of Columbia and two U.S. territories. We found no statement that the project selection would be aimed at particular states or localities because they were hardest hit by the recession.[31]

For GSA projects, all decisions are made at the federal level; therefore, we do not have to consider potential endogeneity introduced by state government level allocation decisions.

Department of Energy (DOE). The Recovery Act authorized $15.55 billion for 10 distinct Energy Efficiency and Renewable Energy (EERE) programs. According to U.S. Dept. of Energy (2009), EERE projects "will stimulate economic development, provide opportunities for new jobs in growing industries, and lay the foundation for a clean energy future." Moreover, "Over $11 billion of EERE's Recovery Act funds will be used to weatherize homes of low-income Americans through the Weatherization Assistance Program (WAP) and will go to states and local communities through the State Energy Program (SEP) and Energy Efficiency and Conservation Block Grant Program (EECBG) to implement high priority energy efficiency projects."

The Recovery Act weatherization component, the largest of the EERE Recovery Act programs, totalled $4.98 billion and were an add-on to the regular annual federal WAP. The Weatherization program state-by-state allocation formula is based on several factors: the low income population, climatic conditions and residential energy expenditures by low income households.

The Department of Energy EERE guidances concerning the Recovery Act do not discuss how states and localities should spend dollars in order to maximize support for areas hardest hit by the recession.[32]

[31] The GSA documents analyzed were General Services Administration (2009a), General Services Administration (2009b) and General Services Administration (2012).

[32] See U.S. Dept. of Energy (2009) and U.S. Dept. of Energy (2010).

U.S. Army Corps of Engineers (USACE) Civil Financing Only program. The $4.6 billion allocated to the USACE was primarily comprised of two parts: Construction ($2 billion) and Operations and Maintenance ($2.075 billion). The spending was applied to improve categories such as inland and coastal navigation, environmental and flood risk management, hydropower and recreation. Besides general provisions applied to all components of Recovery Act funding, the Corps applied the following five additional criteria for project selection: (1) Be obligated quickly; (2) Result in high, immediate employment; (3) Have little schedule risk; (4) Be executed by contract or direct hire of temporary labor; and (5) Complete a project phase, a project, an element, or will provide a useful service that does not require additional funding (see U.S. Army Corps of Engineers (2010a)).

In the agency Recovery Act plans we examined, U.S. Army Corps of Engineers (2010a) and U.S. Army Corps of Engineers (2010b), there was little discussion of the USACE aiming funds towards areas that faced greater economic stress during the past recession. The only exception is that these planning documents mentioned in several places the USACE's desire to "support the overall purpose of ARRA to preserve and create jobs and promote economic recovery; to assist those impacted by the recession; and to provide investments needed to increase economic efficiency." Otherwise, there was no discussion of the USACE aiming targeting project funds to the worst hit areas. Also, there was no specific discussion of how the desire to assist those most impacted by the recession was operationalized in the USACE's plans. Finally, all USACE project decisions are made at the federal level; therefore, there is no potential endogeneity introduced by state government level allocation decisions.

Federal Transit Administration (FTA) Transit Capital Assistance (TCA). The act's Transit Capital Assistance component authorized $6.9 billion in funding for public transit capital improvement, including money allocated to, for example, bus purchases and retrofitting, bus shelters, track rehabilitation and rail cars. Roughly $6 billion of these funds were channeled directly to urbanized areas (UZAs) from the federal government based on apportionment formula. According to Federal Register (2009), "For UZAs with 50,000 to 199,999 in population, the formula is based solely on population and population density. For UZAs with populations of 200,000 and more, the formula is based on a combination of bus revenue vehicle miles, bus passenger miles, fixed guideway revenue vehicle miles, and fixed guideway route miles, as well as population and population density." Note that the amount of aid to each urbanized zone was not dependent on the degree of economic stress felt in the area. Nonurbanized area TCA accounts for $0.68 billion of the program. These grants are made to the state governments and the apportionment formulas are computed based on the ratio of each state's nonurbanized population relative to the national urbanized population as well as the land mass of nonurbanized areas.

Since state governments are the applicants for the nonurbanized area funds, this introduces the potential for endogeneity bias of project selection at the state level. Note that there are no instructions for states to prefer locating projects in areas hit harder by the recession. Whether states themselves took it upon themselves to allocate the nonurbanized area funds to places hardest hit by the recession was not possible for us to discern from available documents. We do note that the nonurbanized program constitutes only a small part of the TCA.

U.S. Department of Education Special Education Fund. The act authorized the Office of Special Education and Rehabilitation Services to allocate \$12.2 billion to states to assist local education agencies in providing free and appropriate public education (FAPE) to students with special needs.[33]

The lion's share of these grant monies came in the form of add-ons to the regular Individuals with Disabilities Education Act (IDEA) Part B funding. The Recovery Act funding formula follow the IDEA Part B formula.[34] The national FFY2009 regular grant amount was \$11.5 billion. The first \$3.1 billion (both from regular funding and the Recovery Act add-on) was divided amongst states so that they were guaranteed to receive their FFY1999 awards. The remaining part of the national award was allocated among the states according to the following rule: "85% are allocated to States on the basis of their relative populations of children aged 3 through 21 who are the same age as children with disabilities for whom the State ensures the availability of a free appropriate public education (FAPE) and 15% on the relative populations of children of those ages who are living in poverty."[35] The Recovery Act add-on totaled \$11.3 billion. Since, at the margin, the FY1999 requirements had already been met by the regular awards, every Recovery Act dollar was in effect assigned according to the 85/15 percent rule.

Next and importantly, we address how funds were assigned from state education agencies to local education agencies (LEA). These initial allocations too were made at the federal level. Each LEA was first allocated a minimum of its FFY1999 award.[36] Beyond these minimums, which were already met by the regular annual award amounts, a slightly different 85/15 rule was used. Within each state, 85% of dollars was allocated to according to the share of school age children in the LEA and 15% was allocated according the LEA's childhood poverty rate. After this, states were allowed to do reallocations as explained below. Before we explain how reallocations worked, we ask whether the observed spending data at the within state level is explained by the simple formulary rules.

Let $P_{j,s}$ and $\tilde{P}_{j,s}$ be the enrollment of students and students in poverty, respectively, in district j and state s. Let $IDEA_{j,s}$ denote the total Recovery Act special needs funding in district j in

[33] Our discussion of the instrument here follows Dupor and Mehkari (2015), who use the special education funding component of the act as an instrument to assess the effect on school districts' spending of the Recovery Act grants.

[34] See U.S. Dept. of Education (2009b) and New America Foundation (2014).

[35] Enclosure B of U.S. Dept. of Education (2009b) contains the precise description of how Recovery Act funds were allocated across states.

[36] Federal code also describes how minimum awards are determined for LEAs created after 1999.

state s. Based on the above formula, the distribution of Recovery Act IDEA dollars would be

$$IDEA_{j,s} = \left(0.85 \times \frac{P_{j,s}}{\sum_{i=1}^{N_s} P_{i,s}} + 0.15 \times \frac{\tilde{P}_{j,s}}{\sum_{i=1}^{N_s} \tilde{P}_{i,s}}\right) IDEA_s$$

Letting P_s and $\tilde{P}_s$ denote the sum within state s of the two district level enrollment variables, we can rewrite this above equation as:

$$\frac{IDEA_{j,s}}{P_{j,s}} = \left[0.85 \times \frac{1}{P_s} + 0.15 \times \frac{1}{\tilde{P}_s}\left(\frac{\tilde{P}_{j,s}}{P_{j,s}}\right)\right] IDEA_s$$

Thus, within each state, the district level per pupil IDEA amount would be perfectly predicted by the ratio of the low-income enrollment to the overall enrollment in the district. By running state-level regressions (available on request) we show that this variable has very little predictive power for the IDEA per pupil amount. This tells us that other factors besides poverty rate in each district are influencing the allocation of IDEA funds.

This brings us to the rules for redistribution of dollars within state across LEAs, given by Code of Federal Regulation 300.707(c)(1). It states:

> If an SEA determines that an LEA is adequately providing FAPE to all children with disabilities residing in the area served by that agency with State and local funds, the SEA may reallocate any portion of the funds under this part ... to other LEAs in the State that not adequately providing special education and related services to all children with disabilities residing in the area served by those LEAs.

We conclude that the primary reason that IDEA money was allocated differently from the formulary rule is that, within individual states, some localities were able to meet their funding requirements of special needs students without using any or all of the Recovery Act IDEA funds. Those funds were then reallocated to districts with additional funding for special needs students. Differences in funding requirements across districts were likely due to various factors, such as the number of special needs students, the types of disabilities and their associated costs and the districts' own funding contributions for providing the services to these special needs students. Our exogeneity assumption is that this set of factors driving redistributions of IDEA funds is orthogonal to the error term in the second stage equation.

A.2 Additional Tables

Table A.1: OLS estimates of the effect on employment of Recovery Act spending, aggregate results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	16.33***	13.45***	12.29***	12.44***	12.27***
($1 million p.c.)	(2.43)	(2.30)	(2.44)	(2.37)	(2.19)
Adjacent ARRA expenditure	4.00***	2.50**	0.52	3.22***	3.28***
($1 Million p.c.)	(1.42)	(0.98)	(0.98)	(1.06)	(0.93)
Job level (2007Q4)	-0.09***	-0.40***	-0.35**	-0.25**	-0.22*
	(0.01)	(0.14)	(0.14)	(0.12)	(0.12)
Income	-	-	-612.63***	-544.77***	-627.31***
(3-yr moving average)†			(168.05)	(145.59)	(142.89)
Log of population†	-	-	0.07	0.11**	0.16**
			(0.04)	(0.06)	(0.06)
Manufacturing share†	-	-	-0.61	-0.76	-1.77**
			(0.46)	(0.46)	(0.87)
Change in the Unemployment	-	-	-0.21***	-0.23***	-0.21***
Rate, Jan. 2008 to Jan. 2009			(0.03)	(0.04)	(0.04)
Proportion of Employment in	-	-	-	-	1.38**
Tradable Sector					(0.68)
Log Change in FHFA HPI, 2002Q4	-	-	-	-	-2.21***
to 2005Q4					(0.48)
Log Change in FHFA HPI, 2005Q4	-	-	-	-	-1.53***
to 2009Q4					(0.38)
Added Employment Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Employment Lags	No	No	No	Yes	Yes
N	1630	1630	1601	1601	1601
R^2	0.340	0.491	0.546	0.576	0.586

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
†Coefficients/SEs are rescaled by 100 to ease interpretation.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.2: First stage least squares estimates of the effect of the composite Recovery Act spending upon direct spending, employment model of direct and spillover results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Composite Instrument	1.56***	1.53***	1.49***	1.45***	1.44***
expenditure ($1 Million p.c.)	(0.16)	(0.16)	(0.15)	(0.14)	(0.13)
Adjacent Composite Instrument	0.13*	0.11	0.09	0.07	0.06
expenditure ($1 Million p.c.)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
Job level (2007Q4)	0.00***	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Added Employment Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Employment Lags	No	No	No	Yes	Yes
Non Labor Controls	No	No	Yes	Yes	Yes
Extra Controls	No	No	No	No	Yes
N	1630	1630	1601	1601	1601
R^2	0.313	0.321	0.323	0.339	0.343
Kleibergen-Paap F-statistic	55.465	56.307	54.205	63.977	62.220

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.3: First stage least squares estimates of the effect of the composite Recovery Act spending upon spillover spending, employment model of direct and spillover results for LM_{1293}

	Pre Recession Level	All Trend Controls	Add Labor Market Controls	Add Region and Spillover Trend Controls (Benchmark)	Extra Controls
	(1)	(2)	(3)	(4)	(5)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Composite Instrument	0.17**	0.15*	0.15*	0.09	0.09
expenditure ($1 Million p.c.)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)
Adjacent Composite Instrument	1.75***	1.74***	1.69***	1.47***	1.46***
expenditure ($1 Million p.c.)	(0.12)	(0.13)	(0.14)	(0.11)	(0.11)
Job level (2007Q4)	-0.00***	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Added Employment Lags	No	Yes	Yes	Yes	Yes
Census Region Dummies	No	No	No	Yes	Yes
Spillover Employment Lags	No	No	No	Yes	Yes
Non Labor Controls	No	No	Yes	Yes	Yes
Extra Controls	No	No	No	No	Yes
N	1630	1630	1601	1601	1601
R^2	0.278	0.280	0.306	0.358	0.362
Kleibergen-Paap F-statistic	55.465	56.307	54.205	63.977	62.220

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.4: Two-stage least squares estimates of the employment response along different degrees of aggregation

	County Level	LM_{1293}	LM_{601}	LM_{372}
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	8.82***	19.30***	22.79***	21.37***
($1 million p.c.)	(3.26)	(5.06)	(7.86)	(6.81)
All Controls	Yes	Yes	Yes	Yes
N	1587	918	510	330
Kleibergen-Paap F-statistic	27.678	118.820	36.115	24.244

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.5: Two-stage least squares estimates of the employment response for various specifications

	Benchmark	Weighted	LM_{372}	LM_{372} $Weighted$
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	10.26***	12.66***	11.49**	16.05***
($1 million p.c.)	(3.84)	(4.10)	(5.34)	(5.80)
Adjacent ARRA expenditure	8.50***	0.53	9.11**	0.62
($1 Million p.c.)	(2.81)	(3.58)	(4.39)	(4.41)
All Controls	Yes	Yes	Yes	Yes
N	1601	1601	655	655
Kleibergen-Paap F-statistic	63.977	31.503	29.966	28.622

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.6: Two-stage least squares estimates of the wage bill and employment response under alternate local ARRA spending definitions.

	Wage Bill		Job-Years	
	ARRA Expenditures	ARRA Obligations	ARRA Expenditure	ARRA Obligations
	(1)	(2)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	0.64***	-	10.26***	-
($1 million p.c.)	(0.22)		(3.84)	
Direct ARRA obligations	-	0.43***	-	6.99***
($1 million p.c.)		(0.12)		(1.75)
Adjacent ARRA expenditure	0.50***	-	8.50***	-
($1 Million p.c.)	(0.17)		(2.81)	
Adjacent ARRA obligations	-	0.19**	-	2.87**
($1 Million p.c.)		(0.08)		(1.37)
All Controls	Yes	Yes	Yes	Yes
N	1601	1601	1601	1601
Kleibergen-Paap F-statistic	69.368	38.167	63.977	35.240

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table A.7: Two-stage least squares estimates of the employment response in the services and goods producing sectors, over various treatment/outcome horizons.

	Thru 2010Q4		Thru 2011Q4		Thru 2012Q4	
	Services Sector	Goods Producing Sector	Services Sector	Goods Producing Sector	Services Sector	Goods Producing Sector
	(1)	(2)	(3)	(4)	(3)	(4)
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Direct ARRA expenditure	4.32*	7.01**	4.80*	9.02**	6.05	11.16**
($1 million p.c.)	(2.22)	(2.86)	(2.78)	(3.92)	(3.68)	(5.05)
Adjacent ARRA expenditure	3.40*	3.38	4.20*	1.22	6.89**	-0.33
($1 Million p.c.)	(1.82)	(2.54)	(2.30)	(2.53)	(3.24)	(3.48)
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	1601	1601	1601	1601	1601	1601
Kleibergen-Paap F-statistic	63.977	63.977	52.491	52.491	37.907	37.907

Notes: The regressions above exclude Alaska, where commuting patterns and the economics of regional markets likely substantially differ from those across the rest of the nation. Regional markets with fewer than 25,000 residents are also excluded. Equations estimated with Huber-White robust standard errors (SEs) clustered by state.
* $p < .1$, ** $p < .05$, *** $p < .01$