The American Recovery and Reinvestment Act: Solely A Government Jobs Program?*

Timothy G Conley[†] and Bill Dupor[‡]

April 15, 2013

Abstract

This paper estimates the private and government sector employment effects of American Recovery and Reinvestment Act (ARRA) spending via an instrumental variables strategy. We argue that this aid was effectively fungible and states used it to offset declines in revenue. This enables us to use exogenous variation in states' budget positions to identify the Act's employment effects. We also exploit exogenous variation across states in ARRA highway funding. According to our benchmark estimates, average state and local government employment, during the twenty-four months following the program's inception, was between 156 thousand and 563 thousand persons greater as a result of ARRA spending (90% confidence interval). The corresponding estimate for the private sector ranges from a loss of 182 thousand to a gain of 1.1 million jobs. Our point estimate for the implied cost of creating a job lasting one year is $202 thousand, which is substantially larger than the corresponding estimate from the President's Council of Economic Advisors.

*JEL Code: E65. Keywords: fiscal policy, fiscal stimulus. All opinions expressed and errors are ours alone. We are grateful to Dean Baker, Mark Bils, David Blau, Henning Bohn, Daniel Bunn, Yongsung Chang, Gabriel Chodorow-Reich, Steve Cohen, Jim Donlin, Duane Dupor, Paul Evans, Laura Feiveson, Joel Gratz, Sam Gratz, Christian Hansen, Don Haurin, Doug Hecox, Damian Kozbur, Igor Livshits, Lance Lochner, Jim MacGee, Rabi Mishalani, Matt Shapiro, Jay Shambaugh, Bahi Subhash, Bruce Weinberg, Daniel Wilson, Mark Wolfgram, Sara Vermillion, Matthew Yglesias as well as other government officials and economists for useful conversations and background. Selina Gonzalez provided outstanding research assistance. We are especially grateful to Chris Robinson for assistance with Census data and Matt Lewis for many useful discussions. An earlier version of this paper circulated under the title "The American Recover and Reinvestment Act: Public Sector Jobs Created; Private Sector Jobs Forestalled." First draft: October 2010.

[†]Department of Economics, University of Western Ontario, Canada. Email: tconley3@uwo.ca

[‡]Department of Economics, The Ohio State University, USA. Email: dupor.1@osu.edu

1 Introduction

As a response to a recession that began in December of 2007, President Barack Obama signed into law The American Recovery and Reinvestment Act, hereafter ARRA, (Public Law 111-5) in February of 2009. It authorized $288 billion for federal tax cuts and $499 billion in Federal Government spending.[1,2] This paper uses variation across states to estimate the causal effect of ARRA spending upon employment, via an instrumental variables strategy.[3]

Our benchmark estimates imply that average employment in the government sector during the 24 months after the passage of the ARRA was 156 thousand to 563 thousand jobs greater as a result of the ARRA spending component. The analogous interval estimate is 82 thousand to 1.55 million jobs for total employment, and between a loss of 182 thousand to a gain of 1.1 million jobs for the private sector. These ranges correspond to 90% confidence intervals and are illustrated in Figure 1 by intervals with circle endpoints.[4]

FIGURE 1 GOES ABOUT HERE

The ARRA spending component was largely a stimulus to state and local governments. The majority of funds were channeled through and controlled by state and local governments and most directly funded jobs were in government (Dupor (2012)).[5] Channeling ARRA aid through state and local government created the possibility that states received different ARRA allocations due in part to differing exogenous capacities to channel or attract federal

[1] Throughout the paper, we use the term spending to include direct government purchases, transfers to state and local governments and entitlements channeled through state and local governments.

[2] The estimated total budget impact of the ARRA, according to the Congressional Budget Office, was increased from its initial $787 billion to $821 billion.

[3] Section 5 discusses other researchers' estimates of the jobs effects of the ARRA in the context of our findings.

[4] In an earlier circulated version of our paper (May 2011), we found that the ARRA had a negative jobs effect in some private sectors. The change in our results seen in the present version arose from a Bureau of Labor Statistics (BLS) data revision.

[5] This included, for example, $86.6 billion to support states' Medicaid programs, a $53.6 billion 'State Fiscal Stabilization Fund' to (in part) aid local school districts, $48.1 billion for transportation infrastructure investment, $40 billion for states to pay unemployment benefits, $13 billion for programs supporting public schools with students from low income families, and $6 billion for clean water projects.

funding. Approximately two-thirds of all ARRA spending was formulary, and there was substantial exogenous state-level variation in formula parameters.

Our benchmark specifications use this exogenous variation in capacity to attract/channel ARRA funds. We exploit variation in funds allocated via the Department of Transportation (DOT), most of which were allocated via formula set prior to the ARRA. The largest component of this, $27.5 billion, was allocated for highway improvements. These dollars were allocated by formulary rule to states, based on pre-defined factors, mainly highway-lane miles, highway usage, and each state's previous contribution to the Federal Highway Fund. This formula was set several years prior to the ARRA's passage and was used to disperse previous highway funds. These outlays should be uncorrelated with each state's short-run budget and general economic situation. Our instrument is a major component of DOT funding relative to pre-recession state revenue.

We also exploit another source of exogenous variation, motivated by the fact that channeling ARRA funds through state and local government created an environment where federal dollars might be used to replace state and local spending. A substantial component of the Recovery Act was authorized specifically to cover states' tax losses (through the State Fiscal Stabilization Fund) and the most dramatic cost increases (through support for state Medicaid programs). The Act did legislate substantial ARRA funds to go to state and local governments for specific programs, such as schools in high poverty neighborhoods and highway construction. However, as depicted in Table 1, states and local governments were already spending significant amounts of their own dollars on many of these programs before the ARRA. Often state spending was substantially higher than nominally targeted ARRA funding. Upon acquisition of ARRA funds for a specific purpose, some states could cut their own expenditures on that purpose. As a result, these states could treat substantial portions of ARRA dollars as general revenue, i.e. these funds were effectively fungible.[6]

[6]We present evidence for such fungibility in the next section. Also, in Section 2, we discuss economic the-

[Table 1 about here.]

Recovery Act aid arrived when many state and local governments were in the midst of budget crises caused by a combination of their reduced tax revenue and increases in expenditures on recession-sensitive programs. These budget pressures provided a substantial incentive for state and local governments to exploit potentially fungible federal aid to pay for their routine operations.[7]

A key assumption in our benchmark specifications is that ARRA dollars were effectively fungible and that states spent ARRA dollars to offset lost revenue. We refer to this as our 'fungibility assumption.' Suppose California loses one dollar in sales tax revenue. If at the same time, California receives an additional ARRA dollar and that ARRA dollar is fungible, then we assume California spends the aid dollar for the same purpose it would have spent its just lost tax dollar. Under this scenario, the relevant treatment is ARRA funding net of state budget shortfalls. This allows us to use exogenous variation in budget shortfalls to identify the effect of ARRA spending. This provides a substantial increase in precision relative to estimates using only exogenous variation in ARRA funding across states.

We construct an instrument to reflect the exogenous component of state budget stress: the pre-recession fraction of revenue that came from inelastic sources. A state that relies on revenue from inelastic sources will experience less fiscal stress during a recession than a state that relies on relatively elastic sources. The relative importance of elastic versus inelastic revenue sources for a state is largely determined by long-run political economy factors which are plausibly unrelated to short run shocks associated with a recession. Our main inelastic sources are property taxes, revenue from publicly-run enterprizes such as universities and liquor stores, as well as sales taxes on cigarettes and alcohol. Elastic sources are mainly

ory, beginning with Bradford and Oates (1971), and empirical work that followed, on how local governments' spending changes with receipt of federal grants.

[7]Aizenman and Pasricha (2011) use aggregate data to show how in large part the ARRA federal expenditure only offset state fiscal declines.

income and general sales taxes.

We estimate sector-specific employment impacts of ARRA spending for a partitioning of non-farm employment into government, goods-producing, and service sectors. The latter two sectors partition the private sector. We think allowing different parameters and using sector-specific conditioning information across goods-producing and service sectors is critical due to their very different pre-recession trajectories and because recessions tend to have a more pronounced impact on markets for goods than services.

The two most commonly reported estimates of the ARRA's jobs effect are those of the Congressional Budget Office (CBO) and the President's Council of Economic Advisors (CEA).[8] Figure 1 compares our estimates of the average employment effects in the first 24 months after the beginning of ARRA spending to their analogs derived from CBO and CEA estimates. The line with triangle endpoints represents the CBO and the interval with diamond endpoints is the CEA range of estimates for job creation/saving.[9]

Our results substantially differ from both the CEA and CBO ranges, though discrepancies are more pronounced for CEA than CBO. The CEA interval for the total effect is between 1.9 and 2.5 million jobs, which is substantially above our estimates. In contrast, there is some overlap between our interval and that of the CBO. The CBO's range for total job creation is very wide, between 0.5 and 2.5 million jobs.

One potential reason that the CEA and CBO jobs effect estimates differ from ours is that we do not consider the stimulative effect of the ARRA's tax cut component; the CBO and CEA each reflect both spending and tax cut components. However, according to several conventional views of fiscal policy, the effect of tax cuts is likely to be much smaller than the effect of government spending. For example, the neoclassical view holds that deficit-financed tax cuts do not increase aggregate consumption because people respond by saving

[8]Section 5 explains the methodology used by these two organizations.

[9]We note that neither CBO or CEA ranges are statistical confidence intervals. Instead, they are ranges based on different projection methods reflecting varied assumptions but not sampling uncertainty.

their additional income in anticipation of future tax liabilities.[10]

Our paper is part of a broader line of research on examining the effects of fiscal policy on real economic activity using variation on sub-national government spending and deficit policy. Four existing studies use this methodology to estimate the jobs effect of the ARRA: Council of Economic Advisors (2009), Chodorow, et. al. (2011), Feyrer and Sacerdote (2012) and Wilson (2011). We discuss our paper's findings relative to these studies.[11]

The next section provides background necessary to understand the Act and evidence related to our fungibility assumption. Section 3 describes the data and regression specifications and Section 4 presents our empirical results. Section 5 discusses other researchers' estimates in the context of our findings. The final section concludes.

2 Background and Evidence of Fungibility

The American Recovery and Reinvestment Act of 2009 (Public Law 111-5) was enacted on February 17, 2009. The Act contains approximately 175,000 words and makes references to hundreds of existing U.S. codes and existing laws. As such, a comprehensive explication of the Act is beyond the scope of this paper.[12]

A key feature of the Act is that a large fraction of the federal dollars were channeled through state and local governments. The Act specified dollar amounts allocated for various categories and often formula for divvying each categories' dollars across states; however, local and state governments had much latitude regarding when and on what projects ARRA dollars were spent. Moreover, each state and local government maintained substantial control

[10] Germane to this view, Shapiro and Slemrod (2008) surveyed recipients of several recent tax cuts, and found that approximately 20% of households will consume the additional after-tax income.

[11] Other related recent papers include Shoag (2010) and Clemens and Mirian (2011).

[12] Michael Grabell's (2012) book *Money Well Spent?: The Truth Behind the Trillion-Dollar Stimulus, the Biggest Economic Recovery Plan in History*, intended for a general audience, gives an interesting narrative account of the effects of the Stimulus. For an early outline of the plan, see Summers (2008).

over how it spent its own non-federal revenues. This is important because it created an environment where ARRA dollars might be used to replace state and local spending.[13]

Consider a specific section of the Act: highway infrastructure improvement. Title XII of Division A of the Act specifies that $27.5 billion shall be allocated to "restoration, repair, construction and other eligible activities," where the eligible activities are spelled out in a particular pre-existing U.S. code. These dollars are divvied up between states based on pre-defined factors, mainly highway-lane miles, highway usage, and each state's previous contribution to the Federal Highway Fund.[14] This formula was set several years prior to the Act's passage and was used to disperse previous highway funds.

Each state selected highway projects on which to spend its dollars. While Federal Highway Administration (FHWA) approval was required, our reading is that the approval rate has been very high.[15] The Act does dictate that the FHWA should give priority to "projects that are projected for completion within a 3-year time frame." The Act also gives a deadline for when grant applications are due, when dollars must be allocated and when the grant dollars must be spent. Agencies provide some guidelines for potential applicants beyond the language of the legislation, e.g. FHWA (2009).

For some components, federal agencies had additional discretion in allocating amounts. For example, the Act allocates $1.1 billion as grants-in-aid for airports. The Act states: "such funds shall not be subject to apportionment formulas, special apportionment categories or minimum percentages ... the Secretary shall distribute funds provided under this heading as discretionary grants to airports, with priority given to those projects that demonstrate to his satisfaction their ability to be completed within 2 years of enactment of this Act."

[13] As Inman (2010) writes, "States are important 'agents' for federal macro-policy, but agents with their own needs and objectives."

[14] The Act specifies a small number of set asides, e.g. $60 million for forest highways on federal land, $20 million for training as well as Federal Highway Administration overhead costs.

[15] The Act does specify that priority will go to "projects located in economically distressed areas." Often, a phrase that might be open to interpretation, such as "economically distressed," is followed by reference to a specific U.S. Code or Law which defines that phrase.

Each of the twenty-eight federal agencies charged with dispersing ARRA dollars submitted "Weekly Financial and Activity Reports." Agencies reported each specific grand (formulary or discretionary) to the Recovery Accountability and Transparency Board, established by the Act. These reports are posted at Recovery.gov and provide a list of awarded grants, obligated amounts and the total outlays for each grant. Outlays are payments from the U.S. Treasury to the grant recipient as directed by the managing federal agency. Most outlays are reimbursements for expenses incurred by the grant recipient. Our main specifications use obligated amounts to construct our treatment measure; we use actual outlays as an alternative treatment measure.

Evidence for Fungibility

States had substantial capacity to treat ARRA dollars as fungible. Establishing this fungibility is important because it will allow us to use exogenous state-level variation in budgetary stress to infer the effects of ARRA spending on employment—thus providing additional variation besides that from exogenous ARRA spending itself. Having this additional variation is important because state variation in ARRA spending, by itself, delivers relatively imprecise estimates of the jobs effect.

The $88.6 billion Medicaid component of the Act provides one example of how states were given enough discretion to treat dollars as fungible. A Council of Economic Advisors (2009) report states that ARRA Medicaid dollars were "intended to boost the level of discretionary funds available to states and not simply to relieve Medicaid burdens."

Similarly, the State Fiscal Stabilization Fund, a $53.6 billion component of the Act, was by design intended to offset the states' lost revenue. According to CBO (2010), it helped "state and local governments stabilize their budgets by minimizing budgetary cuts in education and other essential government services, such as public safety. Stabilization funds for education distributed under the Recovery Act must first be used to alleviate shortfalls in state support for education to local educational agencies and public institutions of higher

education."

Even when ARRA dollars were *notionally* intended for narrow purposes, states often had ways of making the aid effectively fungible. This was the case with the ARRA highway component. For example, in Texas, ARRA highway dollars arrived and simultaneously the number of Texas highway, bridge and street construction workers declined. Employment in that sector fell from 34,600 workers in May of 2008 to 28,500 workers in May of 2010. Total capital outlay on highways in Texas went from $3.38 billion in fiscal year (FY) 2009 to $2.82 billion in FY 2010. This decrease in state expenditures occurred even though Texas spent $700 million in ARRA highway funds during 2010.[16] Thus, the Texas government *responded* to its receipt of ARRA highway dollars by cutting Texas' own contribution to highway spending, which freed up state dollars for other uses. Georgia provides another example of transportation aid fungibility. Georgia's FY 2010 federally-funded transportation capital expenditures were $580 million greater than its FY 2008 level, due primarily to ARRA funds. On the other hand, Georgia's own state funds for these expenditures were $568 million lower in FY 2010 than in FY 2008. Thus, total transportation capital spending was nearly unchanged following the influx of substantial ARRA transportation aid.

Poten and Poten, a private company that collects, analyzes, and sells information about the asphalt industry, describes the situation clearly: "The lack of demand for asphalt is largely due to constrained public road funding and a weak private and commercial market for the product. Most state and local governments have major budget problems. Federal funds related to the transportation budget and stimulus are a critical source of current road funding, but it hardly makes up for the declines from state and local public funding sources, as well as scant private and commercial demand."[17]

[16] Only $110 million in ARRA funds were spent in 2009. The budget amounts come from Texas Department of Transportation (various years).

[17] This quote appears in the August 16, 2010 issue of *Asphalt Weekly Monitor* in the article headlined "US Asphalt Prices Slide Despite It Being Peak Demand Season."

As further evidence that ARRA highway dollars did not translate into a significant increase in highway construction nationwide, consider the number of rebar workers employed in bridge and road building over the past few years. Rebar is a key material in highway construction. Yet, after the ARRA's implementation, there was only a one percent increase in the number of reinforcing iron and steel rebar workers (SOC 47-2171) in the highway, bridge and construction industry (NAICS 237300). Specifically, these workers numbered 1,870 in May of 2008 and 1,890 in May of 2010.[18]

For some components of the stimulus program, language in the Act does try explicitly to prohibit states from cutting state funding upon the receipt of ARRA dollars. U.S. General Accountability Office (2009) states that $101 billion of the spending, including funds for transportation, education and housing, have such restrictions in place. For example, each state governor was required to certify, by March 19, 2009, to the USDOT that the state would maintain a certain contribution to its spending in an area as a condition of accepting ARRA transportation funds. This is part of a "maintenance of effort" requirement of the Act. However, governors were not required maintain pre-Act levels of spending. Rather, it was acceptable to promise to spend less than their expenditure in recent years if a governor could justify the reduction based on other fiscal considerations, such as falling tax revenues.[19]

The use of matching grants, a potential tool to discourage crowding-out, is almost entirely absent from the Act. The section on highway infrastructure investment in the Act states "the federal share payable on account of any project or activity carried out shall be, at the option of the recipient, up to 100 percent of the total cost thereof."

There are however a few components of the ARRA where aid funds were not very fungible. For example, the Act authorized $6 billion for clean-up of nuclear waste sites. Since states spent very little on nuclear clean-up before the ARRA, those state governments could not

[18]BLS 2008, 2010 National Industry-Specific Occupational Employment and Wage Estimates.

[19]The issue of maintenance of effort is complicated. The language of the Act and administrative guidances that followed appear to have interpretations that differ across federal agencies.

free up state dollars by cutting their own spending in response to this $6 billion.

Prior to the Act, researchers had studied whether federal grants crowd out state and local spending. Early analysis by Bradford and Oates (1971) shows that crowding out occurs in a simple political economy model. Empirical work that followed, surveyed in Hines and Thaler (1995), finds evidence against crowding out, which has been termed the "flypaper effect." When a higher-level government issues grants for a particular purpose to a lower-level government, then this money sticks like flypaper towards its intended purpose, with little or no reduction in the lower-level government's contribution. Knight (2002) shows, on the other hand, that after controlling for a particular form of endogenity, the flypaper effect is statistically insignificant in federal highway aid from 1983 through 1997. In the context of our paper, it is worth noting that we are unaware of any studies on the flypaper effect when state and local governments have been under the tremendous budget pressure similar to that observed during the most recent recession.

3 Econometric Specification

We estimate linear models of employment growth as a function of state budget loss, ARRA aid and conditioning variables.[20] Our employment outcome data are derived from monthly series of seasonally-adjusted payroll employment from the BLS Establishment Survey. We use the notation $Y_{j,t}^{Sector}$ for employment in a sector in state j in month t.[21] Our outcome variable for this sector is the average growth in employment from the base of March 2009 over the 24 months from April 2009 to March 2011.

$$\text{Employ Growth}_j^{Sector} = \frac{1}{24} \sum_{k=\text{ Apr 09}}^{\text{Mar 11}} \frac{Y_{j,k}^{Sector} - Y_{j,Mar\ 09}^{Sector}}{Y_{j,Mar\ 09}^{Sector}}$$

[20] A table containing summary statistics for all variables used appears in the online Appendix.

[21] According to our classifications, educators employed by state and local governments count toward the government sector and federal workers count towards service employment.

The sectors we use are: goods-producing, services, and government. We apply our analysis to three distinct employment sectors rather than total employment because of the large differences in trajectories across these sectors over the past decade. Also, state and local government employment fared relatively well during the recession. Employment in the goods-producing sector fell most dramatically during the recession, which was part of a continual decline over the preceding decade.

Time-averaging over a number of months tends to smooth out the idiosyncratic shocks that would be present if we evaluated the ARRA jobs effect at a single particular horizon. This increases our estimates' precision. Moreover, reporting how many jobs were saved/created on average over the two year horizon is a more interesting benchmark–in our view–than reporting how many jobs were saved/created at say a particular (e.g. six month) horizon.

Our first key regressor is AID_j—the ratio of ARRA dollars obligated to state j through March 2011 relative to 2008 state government tax revenue:

$$AID_j = \frac{\text{Total ARRA Obligations April 2009 to March 2011 for State } j}{\text{State } j \text{ Govt Revenue in 2008}}. \tag{3.1}$$

State government revenue includes local government revenue.

Our second key regressor is $LOSS_j$, which measures the decline in state government revenue plus the increase in state Medicaid costs relative to the pre-recession level of government revenue. We include state Medicaid cost increases in this calculation because these increases are mandatory spending and for some states were nearly as substantial as their tax revenue declines. For a state j:

$$\begin{aligned}\text{Net Govt Rev}_j \;\equiv\; &\text{Total Govt Revenue 2009:Q1 to 2011:Q1} \quad -2.25\times\text{Govt Revenue in 2008} \\ &-(\text{Approx. Medicaid Expenditures 2009:Q1 to 2011:Q1}-2.25\times\text{Medicaid Expenditures in 2008}),\end{aligned}$$

Our regressor then normalizes the sign and scales by pre-recession government revenue:

$$LOSS_j = -1 * \frac{\text{Net Govt Rev}_j}{\text{State } j \text{ Govt Revenue in 2008}}. \tag{3.2}$$

A positive $LOSS$ for a state means its budget position has deteriorated. State tax revenue data is from the Census Bureau Quarterly Summary of State Tax Revenue. State-level Medicaid data is from the National Association of State Budget Officers and is calendar-year annual. Details for these data series appear in the Appendix.

We scale ARRA dollars and the decline in states' budget positions by the size of state government revenue rather than the state population because: (i) ARRA funds were channeled largely though the state and local governments that, in turn, used this aid to cover functions that otherwise may have been cut, (ii) the size of government varies substantially across states. Intuitively, even if two states have the same population, each dollar of aid is likely to have a greater impact on a state with a small government relative to a state with a large government. We report results for the alternative, by-population scaling in Section 4.

In our main specifications, we exclude the four least populous states: Wyoming, Vermont, North Dakota, and Alaska. We feel their small populations render their economies sufficiently unusual to warrant their exclusion from our benchmark. In addition, natural resources play an unusually large economic role in three of the four, generating large government revenues. This makes the way in which their state governments operate and respond to ARRA money likely to be substantially different from other states. For example, North Dakota experienced an oil shale boom during the ARRA period.[22] Because of similar timing, the oil-boom related employment growth is potentially confounded with ARRA effects. This further motivates the exclusion of North Dakota from our sample. In our alternate specifications section, we examine the robustness of our results to excluding these four smallest states.

[22]See "Oil Industry Booms–in North Dakota," *Wall Street Journal*, February 26, 2010.

We estimate two main types of regressions for each of three sectors: goods-producing, services, and government. The first type includes the regressors AID_j and $-LOSS_j$ separately. For each sector, our regression specification without our fungibility assumption is:

$$\text{Employ Growth}_j^{Sector} = \alpha_1 AID_j - \alpha_2 LOSS_j + \beta' X_j^{Sector} + \varepsilon_j. \tag{3.3}$$

The second set of regressions impose our fungibility assumption which is equivalent with a coefficient restriction $\alpha_1 = \alpha_2$ in (3.3). Recall that our fungibility assumption is a joint assumption that ARRA funds were fungible and that states used these dollars to offset lost revenue. Under this assumption the relevant treatment variable is aid net of lost revenue: $AID_j - LOSS_j$. Using the notation $OFFSET_j \equiv AID_j - LOSS_j$ we write our benchmark specification as:

$$\text{Employ Growth}_j^{Sector} = \alpha OFFSET_j + \beta' X_j^{Sector} + \varepsilon_j. \tag{3.4}$$

We refer to these specifications (3.4) with fungibility imposed as our benchmarks. In both (3.3) and (3.4) X_j^{Sector} is a vector of state-specific conditioning regressors: the February 2009 employment level in the sector, a region indicator (midwest and northeast), house price growth rates for the run-up period between 2003 and 2006, house price growth rates for the pre-recession decline between 2006 and 2008, 2010 crude oil production per capita, two lags of the annual employment growth in the sector, and a constant.[23]

We rely on X_j^{Sector} to parsimoniously control for states' pre-ARRA economic conditions, which should be predictive of its employment trajectory. The three employment and two house price variables measure these conditions. We add a single region indicator (Midwest/Northeast region) which explains substantial cross-state differences in employment outcomes. This is likely due to the behavior of manufacturing over the cycle combined with

[23]These growth rates are February 2008 to February 2009 and February 2007 to February 2008.

cross-state differences in industrial composition. Finally, we include oil production because oil-extracting states' employment situations and tax revenues can be tied closely to the fuel's price. In our alternate specification section, we experiment with other sets of conditioning information.

Figure 2 contains a scatter plot of states' $LOSS$ versus AID. Note the substantial differences across states on both dimensions. Connecticut (CT) appears in the lower-right corner of the figure. It has a $LOSS$ of 0.325, which means that it experienced a 32.5% decline in its budget position, due to tax losses and Medicaid increases, relative to its 2008 government revenue. It has an AID of 0.115 which means that the ARRA outlays boosted the state's budget by 11.5% relative to its 2008 government revenue. In terms of resources for the Connecticut government to operate, the ARRA dollars were insufficient to cover its losses. Its $OFFSET$ $(= AID - LOSS)$ equals -21%.

South Dakota (SD), on the other hand, fared relatively well during the period, and appears in the upper-left corner. South Dakota has a $LOSS$ of 0.118, which means that it experienced a 11.8% decline in its budget position due to tax losses and Medicaid increases. It has an AID of 0.21, which means that the ARRA boosted the state's budget by 21% relative to its pre-recession size. In terms of resources for the government to operate, the ARRA dollars have more than compensated for the fiscal loss due to the recession in South Dakota. Its $OFFSET$ $(= AID - LOSS)$ equals 0.092 or 9.2%.

Cogan and Taylor (2010) provide evidence that summed across all state and local governments, there was positive accumulation of financial assets during the ARRA period. This could be due to the fact that states in the upper left region of the figure found ways to save (because of fungibility) for precautionary motives rather than spend ARRA aid. This precautionary savings would undo the intended purpose of the ARRA and underscores the need to include states' budget positions in any analysis of the jobs effect of the ARRA.

[Figure 2 about here.]

We estimate equations (3.3) and (3.4) for the goods-producing, services, and government sectors via the generalized method of moments (GMM) using instruments to address the endogeneity of $OFFSET$, AID, and $-LOSS$; treating X as exogenous. The shock to employment growth could reasonably be conjectured to be correlated with the AID and/or $LOSS$ components of $OFFSET$. For example, a negative shock to employment growth in a state might: (i) increase AID if that state received more ARRA dollars because its economy was in worse shape; (ii) increase $LOSS$ if a worse employment situation resulted in lower state tax receipts. We jointly estimate specifications across our three sectors, allowing for error terms to be correlated within state. When constructing standard errors, we impose independence across states and second moment independence of our errors and instruments. Our procedure differs from sector-specific 2SLS only in that we jointly estimate across sectors and allow error terms to be correlated across sectors within a state.

Instruments

Our first instrument is based on dollars allocated by the Federal Department of Transportation (DOT). Roughly $48 billion of ARRA funds were allocated for improving transportation infrastructure, with the largest share going to highway, bridge and intercity rail construction. Much of this was allocated by formula. For example, the Federal Highway Administration (FHWA) was authorized by the ARRA to apportion $27.5 billion. Examples of criteria in the FHWA appropriation formula include each state's: share of total eligible highway lane miles, share of total vehicle miles traveled on eligible highways and share of dollar contributed to the Highway Trust Fund paid by highway users. Our instrument is each state's ARRA dollars from the Federal DOT obligated through September 30, 2010, divided by that state's 2008 tax revenue. We call this our highway spending instrument. The September 2010 end date is motivated by the fact that all formulary FHWA dollars (the largest component of DOT aid) were required to be obligated by this date. After this date, additional obligated dollars were more likely to be non-formulary.

Our second instrument is a ratio of a state's revenue from short-run (relatively) inelastic sources relative to its total revenue, which we refer to as the inelastic revenue instrument. We calculate the ratio for each state using the Census Bureau's 2007-08 Annual Survey of State and Local Government Finances. Using categories defined in this survey, our set of inelastic categories is: property taxes, selective sales taxes, charges and miscellaneous revenue, utility revenue, liquor store revenue, and intergovernmental transfers.

We classify property taxes as relatively inelastic in the short run because most homeowners continue to make their mortgage/insurance/property tax payments even during housing downturns. Also, the downward reassessment of house values, which cause property tax revenues to decline, happens with a substantial lag for most of the housing stock.

"Selective sales taxes" are levied on mainly motor fuel, alcohol, tobacco and public utilities. Demand for these goods tend to be income inelastic. For example, Graham and Glaister (2002) report a short-run income elasticity of auto fuel in the range of 0.35 and 0.55, based on a survey of existing studies.

The "charges" component also responds very little to the business cycle. For example, revenue from government-run universities is a substantial component of the "charges" category, one of the inelastic revenue categories. This revenue growth rate was nearly unchanged in the two years following the start of the recession relative to its growth in the preceding three years. On the other hand, the corresponding decline in state and local individual income tax revenue (not part of inelastic revenue) was over 15%.

Table 2 reports statistics from 'first stage' least squares regressions of our benchmark endogenous variable $OFFSET = (AID - LOSS)$ on instruments as well as the conditioning variables. Also included in the table are results from analogous regressions with AID and $-LOSS$ separately as the outcome. As our logic above suggests, highway spending has statistically significant, positive partial correlations with AID and hence $OFFSET$ with insignificant partial correlation with $LOSS$. Our inelastic-revenue instrument has a

significant, positive correlation with $-LOSS$ which induces a significant, positive partial correlation with $OFFSET$. This instrument does not have a significant partial correlation with AID.

[Table 2 about here.]

The third row in each panel of Table 2 reports partial F-statistics for the joint hypotheses of all instrument coefficients being zero. To evaluate the magnitude of these weak instrument test statistics with $OFFSET$ as the endogenous variable (specification (3.4)), we use critical values provided by Stock and Yogo (2005) based on a researcher's tolerance for a maximal size distortion of nominal 5% level Wald tests involving 2SLS estimators.[24] We view a range of maximal size distortions from 10% to 15% as tolerable; the corresponding Stock and Yogo (2005) critical values are 11.59 and 8.75. Our weak instrument test statistics are large enough relative to this critical value range that we proceed using strong instrument approximations.

Weak instrument (Cragg-Donald) test statistics for specification (3.3) with both AID and $LOSS$ as endogenous variables are 10.26 (Goods), 11.87 (Services), and 8.92 (Government). As above, for evaluating these statistics we use results from Stock and Yogo (2005) for 2SLS weak instrument tests. The Stock and Yogo (2005) critical value corresponding to a 5% maximal size distortion of nominal 5% level Wald tests involving treatment parameters is 7.03. Again, our weak instrument test statistics are large enough relative to this critical value that we proceed using strong instrument approximations.

4 The Act's Jobs Impact

Benchmark Estimates

Table 3 reports α estimates resulting from GMM estimation of (3.3) and (3.4) for goods-

[24] We use 2SLS critical values from Stock and Yogo (2005) because our procedure is close to 2SLS, differing only in that we allow for regression error terms within state to be correlated across sectors and jointly estimate across sectors.

producing, services, and government sectors. In all specifications, X is treated as exogenous and our highway spending and inelastic revenue instruments are used to address endogeneity of $OFFSET$ when fungibility is imposed or endogeneity of both AID and $LOSS$ in the no-fungibility case. Ordinary least squares estimates of the treatment effect parameters are also presented for comparison.[25]

[Table 3 about here.]

The panel labeled "Fungibility not imposed" presents our estimates of the coefficients on AID and $-LOSS$ for each of our three sectors. As anticipated, the signs of the point estimates are positive. There is however a decided lack of precision in these estimates. The only one of the six estimates that is statistically different from zero at conventional significance levels is the coefficient on $-LOSS$ for the government sector, with the AID coefficient perhaps marginally statistically significant in the services sector. In the panel labeled "Fungibility imposed" we present estimates under our fungibility restriction. Our fungibility restriction greatly improves the precision of our estimates relative to those without fungibility. We find a statistically significant, positive effect of $OFFSET$ for the government sector. This is driven in large part by the strong relationship between $-LOSS$ and government sector employment growth that is evident in our results without fungibility. This is a prime illustration of the benefit of our fungibility assumption enabling us to identify the effects of stimulus funds via variation in states' budget stress. The estimated effects of $OFFSET$ in other sectors are substantially smaller and not statistically different from zero at conventional significance levels. The data do not reject the restriction implied by our fungibility assumption, which is unsurprising given the lack of precision in our no-fungibility estimates.[26] Our primary justification for the fungibility assumption remains the narrative evidence for it being a plausible description of how states treated stimulus funds, rather than

[25] See the Appendix for full sets of estimates.

[26] Sector-specific Wald tests of the hypothesis that $\alpha_1 = \alpha_2$ have p-values of .482 (Goods Producing), 0.621 (Services), and 0.606 (Government). The p-value for the analogous joint test across all sectors is 0.528.

the statistical test of its restriction.

Table 4 reports our estimates in terms of numbers of jobs resulting from ARRA spending for total, private-sector and our three individual sectors. These estimates simply combine our treatment parameter estimates in Table 3 with the data on AID and initial employment for each sector.[27]

Our benchmark estimates of jobs effects are reported in the top row of Table 4 labeled "Fungibility imposed." The table reports estimates of the ARRA spending effect upon average employment difference in the 24 months after the implementation of the ARRA relative to February 2009, expressed as thousands of jobs. The pair of numbers beneath each point estimate correspond to a 90% confidence interval.

[Table 4 about here.]

In the state and local government sector, our point estimate is that average employment from April 2009 to March 2011 was 359 thousand persons greater than it would have been in absence of the Act. A 90% confidence interval for this effect is between approximately 156 and 563 thousand. This estimate has a sensible explanation: in states that received relatively less ARRA aid and/or had greater budget losses, government employment was cut or did not significantly expand. In the counterfactual world without the Act, all states would have been forced to take the same action of firing and not filling job openings—resulting in a substantial number of government jobs lost.

In the private sector, the corresponding point estimate jobs effect is positive 456 thousand persons, with a 90% confidence interval of *negative* 182 thousand to positive 1.1 million jobs. Thus our private sector jobs effect has a substantially wider confidence interval than our

[27]The estimated job creation for a sector-state is the product of the sector's estimate of α from (3.3), or α_1 from (3.4) and that states AID times its initial employment in the sector. Thus the aggregate sector-specific jobs estimate is the product of the relevant estimate of α or α_1 and the sum across states of AID times their initial employment in the sector. These cross-state sums of AID times initial employment are 2494 (goods), 11741 (services), 2506 (government) in thousands. Private-sector job estimates are computed via aggregating separate estimates for services and goods-producing sectors. Total employment estimates are similarly computed via aggregating across government, goods-producing and service sectors.

government estimates and it easily includes zero. Aggregating our estimates across sectors results in a point estimate of average employment from April 2009 to March 2011 being 816 thousand jobs greater due to ARRA funds. The corresponding 90% confidence interval is 82 thousand to 1.55 million jobs.

Next, we report the effect of the ARRA in units of its dollar cost per job. Total obligated dollars for the Recovery Act components included in our study was $329 billion (through the Act's first two years across the forty-six states we study). Since average total employment was 816 thousand persons greater for two years (according to our benchmark point estimate), the total job-years employment effect equals 1.63 million. Thus, our point estimate of the cost per job-year equals $202 thousand (= $329B/1.63M).

The second row of Table 4 reports estimated jobs effects without our fungibility assumption. There is a marked decrease in precision without fungibility. For example, the 90% confidence interval for total employment is *negative* 194 thousand to 2.7 million jobs versus an analogous interval estimate of positive 82 thousand to 1.55 million with fungibility. Similar reductions in precision occur for each sector when fungibility is not imposed. While there is a 637 thousand job change in the private-sector point estimates, this is not a large difference relative to the standard errors of the no-fungibility estimates. The decrease in our government sector point estimates of approximately 200 thousand is also substantial but not in our view a large discrepancy given the large no-fungibility standard errors. Thus we consider these intervals' widening when fungibility is relaxed a more salient feature than their shifts.

There are at least two potential reasons that may be behind our low estimates of private sector job creation: (*i*) ARRA funds being focused on government rather than private sector spending not producing large numbers of private sector jobs or (*ii*) there may have been a substantial private sector job creation that was partially offset by crowding out due to government hiring/retention.

Evidence of reason (*i*) was described in detail in Sections 1 and 2. Over one-hundred billion dollars of stimulus spending was targeted directly at the government sector rather than the private sector, e.g. the State Fiscal Stimulus Fund, grants to state and local law enforcement, and support for public schools with students from low income families. These dollars were used largely to pay salaries to avoid government worker layoffs. Moreover, as we have argued above, there is evidence that substantial amounts of aid not directly targeted to the government sector remained fungible, allowing state governments to spend it also in the government sector.

The crowding-out effect, (*ii*), is also plausible as a partial explanation. A portion of the growth in government employment from the ARRA's government-stimulus component was sourced from a pool of potential workers with skills and from regions where there was still private-sector demand. Government workers tend to be well educated. In 2006, the most recent available data, 49% of state and 47% of local government workers had at least a bachelor's degree;[28] for private-sector workers, this proportion is only 25%. The labor market for well-educated individuals was *relatively* strong during and after the recession. In June of 2011, the unemployment rate for the college-educated was 4.4% versus 9.2% for high school graduates without college degrees.[29] By creating/saving government jobs, the ARRA likely prevented some private-sector job creation. The size of this portion remains an open question and the answer is crucial for understanding the Act's jobs effects.[30]

Even apart from government employment crowding out, to the extent that a private-sector job *directly* funded by ARRA spending was filled by an in-demand worker, there may be *zero net private employment increase* from this ARRA job. The Jones and Rothschild

[28]Greenfield (2010).

[29]Both rates apply to high school graduates 25 years or older. The former number applies to those with at least a bachelor's degree and the latter applies to those without bachelor degrees (Bureau of Labor Statistics (2011)).

[30]The magnitude of this type of crowding-out is closely related to the labor market 'mismatch' issue which has been raised, for example, by Federal Reserve Bank of Minneapolis President Narayana Kotcherlakota in several speeches.

(2011a, 2011b) survey finding that approximately one-half of the individuals filling positions directly created by the ARRA were leaving other jobs, is germane to this point.

Estimates Using Alternative Specifications

Table 5 reports jobs effect confidence intervals for alternative specifications. Its second column presents results when an additional instrument is used: an indicator of whether a state had a Democrat as governor prior to the recession. Political considerations may have influenced the enthusiasm with which states pursued ARRA funds. There are several well-publicized instances where Republican governors, including Perry of Texas and Sanford of South Carolina, initially refused to accept parts of ARRA aid.[31] The political party of a state's governor prior to the recession is also arguably a valid instrument as it is predetermined relative to recession shocks. Having a Democrat governor is positively correlated with AID and hence $OFFSET$. Adding this third instrument increases estimation precision but qualitatively these results are similar to our benchmarks.

[**Table 5 about here.**]

The third and fourth columns of Table 5 use alternate definitions for our treatment variable. In the third column, the scaling for AID and $-LOSS$ is per capita rather than scaling by state revenue. In the fourth column we retain our state revenue scaling but use ARRA outlays instead of obligations to calculate AID. Outlays are the amount of ARRA dollars actually paid out versus obligations which include outlays plus amounts committed to the awardees by the respective federal agencies. With our fungiblity restriction imposed, the change to per capita scaling does not have a large impact upon our results. Without this restriction, per capita scaling interval estimates are so wide we view them as essentially uninformative. In contrast, using outlays in place of obligations notably improves our estimates' precision and these shorter intervals are qualitatively consistent with our benchmark results.

The fifth column of Table 5 adds a population regressor to our conditioning informa-

[31] See "They're Saying No to the Cash, but Talk is Cheap" in the Washington Post (2009).

tion. The addition of population to our employment level controls makes the specification better able to capture variation in outcomes related to differential labor force participation across states. A priori we viewed this as potentially important. However, this change in conditioning information does not qualitatively change our benchmark results.

The sixth column of Table 5 explores whether our results are sensitive to a redefinition of *LOSS* to exclude the Medicaid cost increase component. This alternative definition of *LOSS* yields results that are qualitatively very similar to our benchmark results.

Table 6 examines the sensitivity of our results to excluding the four smallest population states. The second column presents results when we estimate our specifications using all 50 states. When our fungibility restriction is imposed, government sector results are similar to our benchmarks. However, there is an appreciable increase in our private sector employment estimates of approximately 300 thousand jobs with little change in precision. A substantial shift upward in private sector interval estimates also occurs without our fungibility restriction but, as with our main results, the estimates are very noisy. Of the four small states, Vermont appears to have the largest individual influence upon our results, illustrated in the third column where it alone is added to our 46 state sample. The analogous results when either Alaska, North Dakota, or Wyoming is added by itself to the sample are similar to the baseline. However, some pairs of these three states move estimates substantially when added to our 46 state sample. The fourth column illustrates a shift up in our no-fungibility results with the addition of only North Dakota and Wyoming to the sample. The fifth column illustrates a shift in our private sector results with the addition of only Alaska and North Dakota. The bottom line is that our main results are sensitive to excluding the four smallest states; however, this sensitivity is not large enough to change the overall tenor of our results.

[Table 6 about here.]

5 Other Researchers' Estimates of Job Creation

Even before the ARRA was passed, Bernstein and Romer (2009) reported that 3.6 million jobs would be created or saved by the then-envisioned legislation, relative to a no-stimulus baseline. This was based on existing estimates of fiscal policy multipliers. Their estimates included both the tax and spending components of the ARRA.

Since the passage of the ARRA there have been two types of studies of its jobs effect. The first type uses *all pre-ARRA data and information except for the actual ARRA spending.* They fail to look at the actual employment outcomes experienced after the program starts. The methodology used in Congressional Budget Office (various quarterly reports) is an example of this pre-ARRA data approach. The CBO estimates that the employment increase attributable to the ARRA was between 0.5 and 2.5 million between April 2009 and March 2011, as shown in Table A.4 in the Appendix.[32] To construct these numbers (in their Table 1), the CBO divides the total spending of the ARRA into its components and then applies low and high output multipliers. These multipliers were delivered from the CBOs analysis of existing studies.

The CEA uses two methodologies to assess the jobs impact of the ARRA. The first of these is a multiplier approach similar to the CBO. As illustrated in Figure 1, the Council of Economic Advisors (various quarterly reports) estimates an effect of increasing employment by 1.9 to 2.5 million workers on average over the Act's first 24 months. This contrasts with our benchmark estimates of an increase of 82 thousand to 1.549 million workers due to ARRA spending. Our results are more consistent with the CBO estimates as our interval estimate does overlap with the CBO range.

The second CEA methodology involves estimating a time series model used to predict employment outcomes for the post ARRA period and uses these predictions as a counter-

[32]Note these ranges are computed based on both government spending and tax cut incentives in the Act.

factual for employment in the absence of stimulus activity. Specifically, the CEA estimates a vector autoregression of employment and output using data from 1990:Q1 to 2007:Q4 to construct an employment forecast from 2009:Q2 to 2010:Q2. They interpret the resulting forecast errors as being due to stimulus policy. According to these estimates (see appendix Table A.4), at the end of 2010:Q2, the Act had increased employment by 3.6 million workers. As the CEA notes, this procedure can only measure the 'net effect' of all combined stimulus components, e.g. fiscal as well as monetary policy actions. Moreover, any stimulus effects are inherently confounded with coincident shocks to the economy, e.g. financial shocks driving European debt concerns. Thus these estimates are not directly comparable to our estimates of the employment effects of ARRA spending.

The second type of studies are more closely related to ours and use employment observed during and after the ARRA period to estimate a jobs effect. We discuss our paper's findings relative to these studies below.

Feyrer and Sacerdote (2012) conduct both cross-sectional and time series analyses to estimate the employment effects of the ARRA. Based on state-level data, their cross-section point estimates imply that each job cost $107,000, while their time series point estimates imply a cost-per-job of $400,000. Feyrer and Sacerdote (2012) use a measure of the seniority of each state's Congressional delegation an an instrument for ARRA spending, although the specific estimates they highlight, given above, are based on least squares.

Wilson (2011) estimates the job effects of ARRA spending using state-level variation via an instrumental variables method. Wilson's overall results suggest a larger jobs effect of ARRA spending than we find, though our results are compatible in some aspects. According to his preferred specification, Wilson's point estimate implies that the average cost-per-job was $125,000. Though we share an econometric method, we use different treatment measures, outcomes, conditioning information, and instruments. Thus, our results provide a complementary look at ARRA job effects.

Wilson's treatment measures differ from $OFFSET$ in that he uses a per capita scaling. In addition, his preferred stimulus measure is funds at an 'announcement' stage months prior to the funds being obligated to specific recipients, constructed from his own careful collection of non-archived federal funding announcements.

Wilson's outcomes are per capita changes in employment from February 2009 to different horizons during the following two years, with February 2010 as his preferred horizon. Our outcome is an average of employment changes from April 2009 to March 2011. We focus on this average effect because it is more precisely estimated than month-by-month effects in our specifications. Wilson's preferred estimates imply a 90% confidence interval estimates of the jobs effect of announced ARRA spending in March 2010 of approximately 600 thousand to 3.5 million jobs. This differs from our benchmark interval estimate for the average jobs effect of 82 thousand to 1.549 million jobs. Leaving aside differences due to announced versus obligated spending, it is possible for the impulse response to spending to be consistent with Wilson's range in March 2010 while still on average being consistent with our range. Our average response measure could be below the level of the response midway simply due to responses in the first few months post ARRA being small.

Wilson's conditioning variables (included exogenous regressors) also differ from ours. We think the most important differences are that he uses less information about lagged employment than we do and includes a measure of the change in personal income. The former is simply a difference in judgement calls regarding a parsimonious characterization of state employment trajectories. The latter is more directly motivated since, in our framework, personal income changes are not a valid exogenous included variable because they are directly related to our endogenous state revenue loss measure, $LOSS$ (and hence $OFFSET$). They are more easily argued to be exogenous in Wilson's framework.

Wilson uses instruments that measure the formulary component of ARRA aid administered through the Departments of Transportation, Education and Health and Human Ser-

vices (HHS). For example, Wilson's Transportation instrument is a linear combination of the three factors that determine states' highway funding: state highway lane miles, vehicle miles traveled and payments to the highway trust fund. This instrument is similar at least in motive to ours. His Education instrument is the school-age share of the state population, which is a key factor in determining ARRA education aid. Wilson's HHS instrument is nearly identical to the formulary component of Medicaid Assistance Program aid used by Chodorow, et. al. (2011), which we discuss below.

Wilson focuses on aggregate specifications while we allow separate parameters by sector when constructing aggregate employment effects. We pay a cost in precision for the benefit of flexibility. However, our benchmark results are compatible with his sector-specific preferred results. He also finds statistically insignificant private sector effects and positive, statistically significant government job effects for his preferred treatment variable of announced ARRA funding through February 2010.

Estimates across studies, generally imply that ARRA job creation was expensive: $202,000 (this paper), $107,000 and $400,000 (Feyrer and Sacerdote (2012) cross-section and time series, respectively), and $125,000 (Wilson (2011)).[33]

The sole study that finds inexpensive job creation is Chodorow, et. al. (2011) who study the employment effect of the Federal Medicaid Assistance Program. This component of the ARRA (approximately 17% of non tax cut ARRA costs) provided state governments with a one-time supplement to the ongoing federal aid for state Medicaid programs. The study's estimate of the cost of creating one job-year is $26,000.

Dupor (2013) shows that two issues may be critical understanding the difference between Chodorow, et. al. (2011) and other studies.[34] First, Chodorow, et. al. (2011) includes

[33]Comparability across studies is somewhat limited because of differences in precise definitions of 'a job' across studies.

[34]A second possible explanation would be that the Medicaid component was more effective than the other parts of the ARRA. We find this explanation uncompelling because of the fungibility of the different types of ARRA aid.

Washington DC in its sample, whereas the other studies include only U.S. states.[35] Washington DC had very strong employment growth relative to other states, largely due to the acyclicality of the federal workforce. Moreover, the instrument used by Chodorow et.al. is the formulary component of ARRA Federal Medicaid assistance and Washington DC is a substantial outlier for this variable. Second, the controls used by Chodorow, et. al. (2011) differ substantially from other studies, particularly in the specification for lagged employment growth, and their omission of house price growth controls.[36] Dupor (2013) repeats the Chodorow, et. al. (2011) analysis except that he drops Washington DC from the sample and with controls modified to more closely align with our study. These adjustments lead to an increase in the point estimates of the cost-per-job (relative to Chodorow et.al.) by roughly an order of magnitude, making it comparable to our results.

6 Conclusion

Our findings suggest that, relative to a no-stimulus baseline, ARRA spending expanded government employment. The private-sector employment effect was statistically not different from zero, and estimated with sufficient precision to rule out a large jobs impact. Our empirical strategy relies on the key role played by state and local governments in implementing this federal stimulus policy. This allowed us two exogenous sources of variation to identify the employment effects of ARRA spending: cross-state variation in budget stress, via our fungibility assumption about how state and local governments used stimulus funds, and variation in a formulary component of ARRA spending.

An indispensable role for state and local governments appears to have been necessary given the Act's objectives. First, putting hundreds of billions of dollars into the economy

[35] We do not include Washington DC in our sample because some of our data are available only for states, the DC government structure is quite different from that of states, and it is also a small population region.

[36] We use a pre-recession employment level and two lags of annual growth, Wilson uses one lag of annual growth. Chodorow, et. al. (2011) use a lagged five-month employment growth regressor.

quickly required the federal government to delegate substantial authority to state and local governments. Second, the drive to spend rapidly together with the goal of strict oversight led these governments to execute stimulus by maintaining and expanding existing public-sector programs. It is perhaps unsurprising, therefore, that the Recovery Act's stimulative effect was mainly in government.

Much work on the effects of the ARRA remains to be done.

First, researchers must employ additional restrictions or data, beyond simply cross-state variation in spending, to improve the precision of jobs effect estimates. Other researchers working with cross state data find wide confidence intervals and we find wide confidence intervals when we do not impose fungibility. Our fungibility restriction substantially improves precision, leading us to think that more explicit economic modeling of state governments' aid spending decisions could yield additional useful restrictions.

Second, researchers should allow for cross-state and region spillovers. This might result in estimates of a larger overall jobs effect. Suppose, for example, that Georgia received relatively more ARRA aid, which in turn stimulated that state's economy. If, as a result, Georgia residents' vacation spending in Florida increased, then the increased vacationing might generate jobs in Florida. Our methodology cannot pick up this effect.[37]

Next, research on the ARRA demands greater structural economic modeling. In this study, we deliberately chose a relatively model-free approach for one of the first studies on this new government program and data set. The drawback is that, at this point, we can only conjecture on the underlying, economic mechanisms that give rise to our findings. A reliable dynamic general equilibrium model of the ARRA experience will likely need to address three issues. First, the most recent recession and the Act's implementation happened partway through a sectoral shift from goods-producing to a service (of a particular type) sector

[37]Another type of cross-state spillover is possible because U.S. states share a common monetary policy (see Nakamura and Steinsson (2010)).

economy.[38] The way that governments spent ARRA dollars across different sectors likely had consequences for the Act's impact. Second, relatively less educated workers faced a much worse job market than more educated workers. For this reason, it will likely be important to differentiate between high and low education workers. Third, we have provided evidence that state and local governments have used part of ARRA aid in a way not explicitly intended by Congress and the President. As such, we think it will be important to include a hierarchal intergovernmental component in the spirit of Bradford and Oates (1971).

References

American Recovery and Reinvestment Act, Public Law 111-5, 111*th* Congress, 17 Feb 2009

Aizenman, J. and G.K. Pasricha (2011), "The Net Fiscal Expenditure Stimulus in the U.S., 2008-2009: Less than What You Might Think, and Less than the Fiscal Stimuli of Most OECD Countries," *The Economists Voice,* Berkeley Electronic Press, June.

Bernstein, J. and C. Romer (2009), "The Job Impact of the American Recovery and Reinvestment Plan," Office of the Vice President-Elect, working paper.

Blanchard, O. and R. Perotti (2010), "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics* 117, 1329-1368.

Bureau of Labor Statistics (2011), "The Employment Situation–June 2011," June 8, 2011.

Blinder, A. and M. Zandi (2010), "How the Great Recession Was Brought to an End."

Bradford, D. and W. Oates (1971), "The Analysis of Revenue Sharing in a New Approach to Collective Fiscal Decisions," *Quarterly Journal of Economics* 85, pp. 416-439

[38]Existing economic modeling of sectoral shifts include Davis and Haltiwanger (2001) and Phelan and Trejos (2000).

Chodorow-Reich, G., L. Feiveson, Z. Liscow and W. Woolston (2011), "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," University of California Berkeley, working paper.

Clemens, J. and S. Miran (2011), "Fiscal Policy Multipliers on Sub-National Government Spending," Stanford University, working paper.

Cogan J. and J. Taylor (2010), "What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package," NBER working paper 16505.

Congressional Budget Office (various quarterly reports), "Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output."

Council of Economic Advisors (various quarterly reports), "The Economic Impact of the American Recovery and Reinvestment Act of 2009."

Council of Economic Advisors (2009), "The Effects of State Fiscal Relief."

Davis, S. and J. Haltiwanger (2001), "Sectoral Job Creation and Job Destruction in Response to Oil Price Shocks," *Journal of Monetary Economics* 48, pp. 465-512.

Dupor, B. (2012), "Directly Created and Saved Jobs Were Predominantly in Government: An Analysis of the ARRA Recipient Reports" The Ohio State University.

Dupor, B. (2013), "Cross-Sectional Analyses of the Employment Effects of the 2009 Recovery: Explaining the Differences across Studies " The Ohio State University.

Federal Highway Administration (2009), "American Recovery and Reinvestment Act of 2009 Implementation Guidelines," U.S. Department of Transportation, April 1 (updated).

Feyrer, J. and B. Sacerdote (2012), "Did the Stimulus Stimulate? Effects of the American Recovery and Reinvestment Act," Dartmoth College.

Graham, D. and S. Glaister (2002), "The Demand for Automobile Fuel: A Survey of Elasticities," *J. of Transport Economics and Policy* 36, pp. 1-25.

Grabell, M. (2012), *Money Well Spent?: The Truth Behind the Trillion-Dollar Stimulus, the Biggest Economic Recovery Plan in History,* Public Affairs.

Greenfield, Stuart (2010), "Public Sector Employment: The Current Situation," Center for State and Local Government Excellence, working paper.

Hines, J. and R. Thaler (1995), "The Flypaper Effect," *J. of Econ Perspectives:* 9, 217-226.

Inman, Robert (2010), "States in Fiscal Distress," NBER Working Paper 16086.

Knight, B. (2002), "Endogenous Federal Grants and Crowd-out of State Government Spending: Theory and Evidence from the Federal Highway Aid Program," *Am. Economic Rev.*, pp. 71-92.

Nakamura, E. and J. Steinsson (2010), "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions,' unpublished manuscript, Columbia University.

National Association of State Budget Officers (various years), State Expenditure Report, various years.

Phelan, C. and A. Trejos (2000), "The Aggregate Effects of Sectoral Reallocation," *Journal of Monetary Economics* 45, pp. 249-268.

The Recovery Accountability and Transparency Board (2009), "Recovery.gov: Download Center Users Guide."

Shoag, D. (2010), "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns,", Harvard University, working paper.

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.

Summers, L. (2008), "Update on Economic Policy Work," memo to President-Elect Obama, December 15.

Texas Department of Transportation (various years), "Annual Financial Report (unaudited)" 2009 & 2010.

U.S. Census Bureau (2009), "2008 Annual Surveys of State and Local Finances."

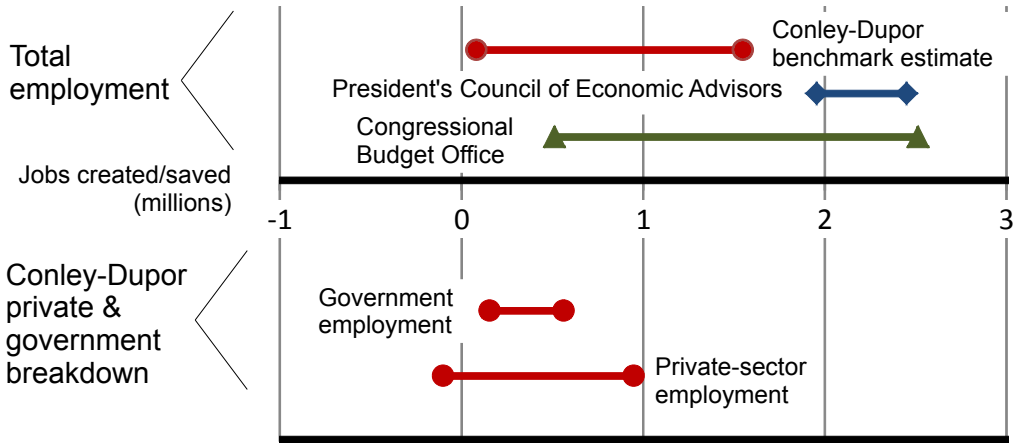
U.S. Department of Social Services (2010), "TANF Emergency Contingency Fund: American Recovery and Reinvestment Act (ARRA) of 2009," PowerPoint presentation, November.

U.S. General Accountability Office (2009), "Planned Efforts and Challenges in Evaluating Compliance with Maintenance of Effort Provisions," GAO-10-247, November.

Washington Post (2009), "They're Saying No to the Cash, but Talk Is Cheap," March 15.

Wilson, D., "Fiscal Spending Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act," Federal Reserve Bank of San Fransico, working paper, 2010.

Figure 1: Our benchmark estimates of employment effect (in millions) of the ARRA spending component, with Congressional Budget Office and President's Council of Economic Advisors ranges shown for comparison



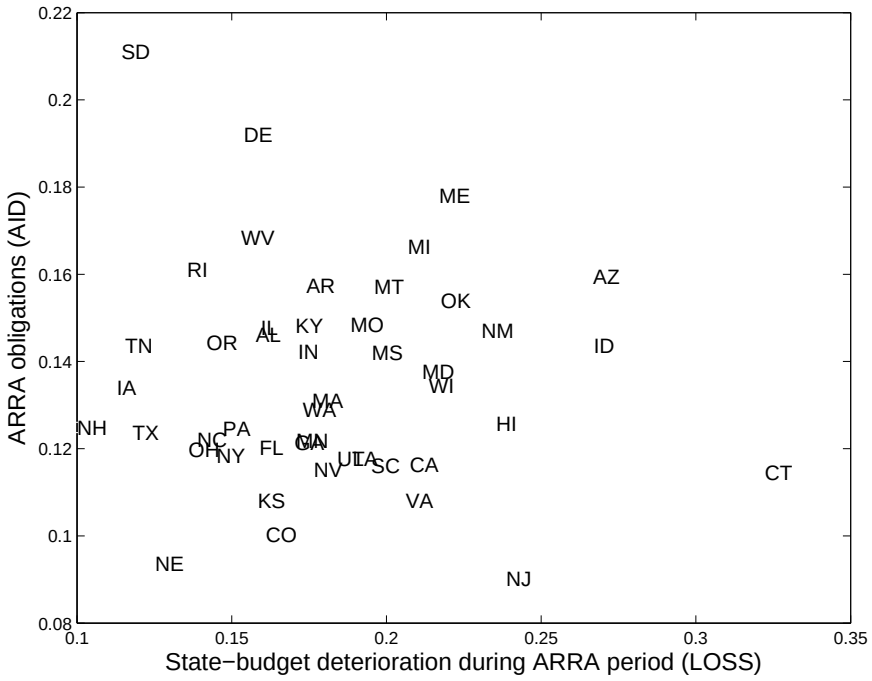
Notes: Benchmark estimates are 90% confidence intervals for employment due to ARRA-spending component (averaged between March 2009 and March 2011). Other sources are Council of Economic Advisors (various quarterly reports) and Congressional Budget Office (various quarterly reports). The CEA and CBO ranges each include the effects of tax cuts and spending. See Table A.4 in the Appendix for data underlying the figure.

Table 1: State government contributed spending and ARRA spending, various categories and in billions of dollars

Spending category	ARRA	States-contributed spending
Medicaid	88.6	271.8
Elementary and secondary education	53.6	464.7
Highways	28.0	50.1

Notes: ARRA dollars are amounts authorized by the Act. State contributions are 2008 FY multiplied by two: the number of years of ARRA spending covered by this study. Elementary and secondary education of ARRA dollars refers to State Fiscal Stabilization Fund. State-contributed education spending does not include $557 billion 2008 spending by local governments. Sources are Public Law 111-5 (2009), National Association of State Budget Officers (various years) and U.S. Census Bureau (2009).

Figure 2: Each state's revenue loss net of Medicaid increase and ARRA obligations, through March 2011, relative to its pre-recession revenue, i.e. $LOSS$ and AID



Notes: Excludes four smallest states.

Table 2: T- and partial F-statistics for instrument relevance

	$OFFSET$ (=$AID - LOSS$)	AID	$-LOSS$
	Goods producing		
T statistics			
Highway spending	2.75	5.23	0.50
Inelast. revenue	3.20	-1.18	4.50
Partial F-statistic	10.31	13.79	10.79
	Services		
T statistics			
Highway spending	2.82	5.10	0.49
Inelast. revenue	3.60	-0.81	4.73
Partial F-statistic	12.10	12.99	11.87
	Government		
T statistics			
Highway spending	2.46	4.63	0.47
Inelast. revenue	3.14	-0.63	4.08
Partial F-statistic	9.32	10.73	8.92

Notes: Above t-statistics are from the single-equation least-squares estimation of alternative endogenous variables on instruments and all control variables. Partial F-statistic reflects joint test against the null that coefficient on every instrument equals zero.

Table 3: Treatment coefficient estimates of response of employment growth (monthly average between April 2009 and March 2011) to Recovery Act aid

(a) Instrumental variables

	Goods-producing	Services	Government
	Fungibility imposed		
AID-LOSS	0.063	0.025	0.143
	(0.091)	(0.022)	(0.049)
	Fungibility not imposed		
AID	0.157	0.060	0.065
	(0.169)	(0.045)	(0.104)
-LOSS	0.028	0.013	0.174
	(0.103)	(0.026)	(0.061)

(b) Least squares

	Goods-producing	Services	Government
	Fungibility imposed		
AID-LOSS	-0.014	0.015	0.103
	(0.053)	(0.014)	(0.028)
	Fungibility not imposed		
AID	0.210	0.022	0.063
	(0.107)	(0.028)	(0.062)
-LOSS	-0.094	0.013	0.116
	(0.061)	(0.016)	(0.034)

Notes: Results labeled 'Fungibility Imposed' correspond to our benchmark specification (3.4) α. Results labeled 'Fungibility Not Imposed' correspond to specification (3.3) (α_1, α_2). IV results use our highway spending and inelastic revenue instruments. Standard errors appear in parenthesis.

Table 4: Employment effects of ARRA spending: monthly average between April 2009 and March 2011, reported as thousands of jobs (ninety percent confidence intervals)

	All sectors	Private	Government	Goods-producing	Services
			Instrumental variables		
Fungibility	816	456	359	158	299
imposed	(82 , 1549)	(-182 , 1095)	(156 , 563)	(-214 , 529)	(-117 , 714)
Fungibility	1256	1093	163	391	702
not imposed	(-194 , 2706)	(-155 , 2340)	(-267 , 594)	(-303 , 1084)	(-159 , 1563)
			Least squares		
Fungibility	401	144	257	-34	179
imposed	(-35 , 838)	(-244 , 532)	(141 , 374)	(-253 , 185)	(-83 , 440)
Fungibility	945	786	159	522	264
not imposed	(39 , 1850)	(-11 , 1582)	(-96 , 414)	(82 , 962)	(-284 , 811)

Notes: Results labeled 'Fungibility Imposed' correspond to our benchmark specification. Using two instruments and 46 states. Ninety-percent confidence intervals appear in parenthesis. With the fungibility restriction, the p-value for the test of overidentifying restrictions is 0.532.

Table 5: Employment effects of ARRA spending: monthly average between April 2009 and March 2011, various state combinations, reported as thousands of jobs (ninety percent confidence intervals)

	Benchmark	Add Dem governor instrument	Use per-capita scaling	Outlays instead of obligations	Add population regressor	LOSS includes tax only
Private-sector						
Fung. imposed	(-182 , 1095)	(-117 , 1034)	(-189 , 1243)	(-162 , 826)	(-150 , 1139)	(-131 , 1144)
Fung. not imposed	(-155 , 2340)	(-107 , 2192)	(-303 , 4275)	(-130 , 2063)	(-96 , 2651)	(-284 , 2306)
Government						
Fung. imposed	(156 , 563)	(172 , 544)	(169 , 659)	(126 , 435)	(153 , 558)	(119 , 517)
Fung. not imposed	(-267 , 594)	(-208 , 579)	(-622 , 798)	(-230 , 514)	(-256 , 592)	(-505 , 393)
Total						
Fung. imposed	(82 , 1549)	(156 , 1478)	(105 , 1778)	(48 , 1177)	(115 , 1585)	(103 , 1546)
Fung. not imposed	(-194 , 2706)	(-106 , 2563)	(-530 , 4678)	(-159 , 2376)	(-111 , 3002)	(-543 , 2453)

Notes: Negative numbers correspond to jobs destroyed/forestalled.

Table 6: Employment effects of ARRA spending: monthly average between April 2009 and March 2011, various state combinations, reported as thousands of jobs

	Benchmark	Use all 50 states	46 plus Vermont	46 plus ND and WY	46 plus AK and ND
Private-sector					
Fung. imposed	(-182 , 1095)	(124 , 1456)	(-86 , 1480)	(-205 , 1312)	(124 , 1073)
Fung. not imposed	(-155 , 2340)	(504 , 2932)	(220 , 2557)	(181 , 2804)	(-28 , 2616)
Government					
Fung. imposed	(156 , 563)	(130 , 444)	(186 , 695)	(165 , 563)	(118 , 384)
Fung. not imposed	(-267 , 594)	(-483 , 500)	(-107 , 745)	(-203 , 607)	(-578 , 409)
Total					
Fung. imposed	(82 , 1549)	(351 , 1803)	(221 , 2054)	(89 , 1746)	(331 , 1368)
Fung. not imposed	(-194 , 2706)	(344 , 3109)	(317 , 3098)	(234 , 3155)	(-248 , 2667)

Notes: Negative numbers correspond to jobs destroyed/forestalled.

An Online Appendix for "The American Recovery and Reinvestment Act: Solely A Government Jobs Program?"

April 2013

This appendix provides supplementary material for "The American Recovery and Reinvestment Act: Solely A Government Jobs Program?" by Timothy G. Conley and Bill Dupor.

A Online Appendix

A.1 Data Construction

Table A.1 provides summary statistics for data used in this paper.

The numerator for AID_j, as defined by equation (3.1) uses ARRA obligated dollars taken from the federal agencies' "Weekly Financial and Activity Reports," which are cumulative over time and available on the recovery.gov web site.[39] The Social Security Administration dollars are excluded entirely because these are direct transfers paid by the Federal Government to citizens. We also exclude the following agencies from our data: General Services Administration, National Science Foundation, Small Business Administration, Department of Interior, Federal Communication Commission, Department of State, Retirement Railroad Board, Department of Veterans Administration, NASA. Each of these exclusions is due to one of the following reasons: data was not reported in a useful manner, none of that agencies' funding was reported as going through the states. The total funding outlays from excluded agencies (expect for the SSA) make up less than 2% of total ARRA outlays.

The denominator for AID_j requires a measure of pre-recession size of state revenue. Because we do not need data since the recession's beginning, we are able to use a more

[39] Documentation for the web site data is contained in Recovery Accountability and Transportation Board (2009)

Table A.1: Summary statistics

Variable	Mean	Stdev.	10th perc.	90th perc.
Economic and financial variables				
Obligated AID	0.136	0.024	0.108	0.161
$LOSS$	0.178	0.045	0.118	0.231
Avg. Goods-producing employment growth (percent)	-7.01	2.70	-9.63	-4.86
Avg. Services employment growth (percent)	-0.66	0.53	-1.33	0.01
Avg. Government employment growth (percent)	-0.71	1.34	-2.59	1.04
2009 Government employment (millions)	0.43	0.42	0.09	0.72
2009 Services employment (millions)	2.00	2.00	0.42	4.15
2009 Goods-producing employment (millions)	0.42	0.40	0.06	0.84
ARRA obligated dollars (billions)	7.179	7.777	1.516	13.670
Instruments				
USDOT-obligated aid scaled by gov. revenue	0.02	0.01	0.01	0.02
Intensity of inelastically-sourced revenue (percent)	75.10	7.00	67.58	82.24
Democratic governor (fraction)	0.59			
Other variables				
2010 oil production (barrels per capita)	3.63	7.48	0.00	14.74
House price growth, 2003 to 2006 (percent)	0.23	0.13	0.10	0.42
House price growth, 2006 to 2008 (percent)	-0.05	0.14	-0.27	0.07
Midwest/Northeast (fraction)	0.43			
Observations	46			

Notes: The job growth data is from non-farm employment. The four lowest population states (i.e. Alaska, North Dakota, Vermont and Wyoming) are excluded from the sample. See text for variable definitions.

accurate measure of the size of state government–one that includes non-tax state revenue sources as well as local governments' revenue. This is the calendar-year state revenue from all sources, which is reported in the "Annual Surveys of State and Local Government Finances."

This measure of pre-recession state government revenue is also the denominator for $LOSS_j$, as defined by equation (3.2). Its numerator requires a measure of total tax revenue collected by state government. This is collected by the U.S. Census in the "Quarterly Summary of State and Local Government Tax Revenue," which is available on the Census web site.[40] In addition to state revenue, we use state Medicaid outlays (independent of the funding source). Medicaid data is available in National Association of State Budget Officers (various years). Unlike the quarterly Census-collected tax revenue data, the state-level NASBO Medicaid data is reported in annual, calendar amounts. Therefore, we are unable to match Medicaid cost to the timing of state tax revenue exactly.

Employment is the non-farm payroll series, de-seasonalized, from the Bureau of Labor Statistics Establishment Survey. This gives the number of workers in each state by month. We use measures from three sectors: state plus local government, goods-producing and services. Each series has been de-seasonalized either by us, using the Census' X12 algorithm, or by the Census itself. Further details on the employment data appear in the body of the paper.

Our measures of inelastic revenue share are from the 2007-08 "Annual Surveys of State and Local Government Finances." In the denominator, we use reported total state government revenue from all sources. Using categories defined in this survey, our set of inelastic categories is: property taxes, selective sales taxes, charges and miscellaneous revenue, utility revenue, liquor store revenue, and intergovernmental transfers. The main revenue categories implicitly defined as elastic are: income taxes, general sales taxes, and insurance trust rev-

[40]Note that state tax is less than actual state government income. This is because we do not have up-to-date data on the two main others sources of state income: earnings from other sources, e.g. university tuition, and non-ARRA Federal aid.

Table A.2: Response of employment growth (monthly average between April 2009 and March 2011) to state government losses and obligated ARRA money, fungibility imposed, all coefficients reported

	Goods-producing	Services	Government
AID-LOSS	0.06330	0.02545	0.14339
	(0.09059)	(0.02151)	(0.04942)
Pre-recession level (in billions)	0.01698	0.00035	-0.00082
	(0.00748)	(0.00037)	(0.00350)
One-year lag employ. growth	0.07855	0.18078	0.05329
	(0.09145)	(0.08097)	(0.12478)
Two-year lag employ. growth	-0.34595	-0.24858	-0.01079
	(0.16146)	(0.13701)	(0.15539)
Midwest/Northeast region	0.02554	-0.00282	0.00045
	(0.00653)	(0.00195)	(0.00353)
Oil-production per capita	0.22684	-0.08748	0.24652
	(0.43185)	(0.10401)	(0.22291)
House price grow (03-06)	-0.01461	0.00019	0.00155
	(0.03056)	(0.00715)	(0.01498)
House price grow (06-08)	0.14633	0.01289	0.03987
	(0.03164)	(0.00846)	(0.01410)
Constant	-0.07357	0.00403	-0.00050
	(0.01240)	(0.00438)	(0.00481)

enue.

The remaining data are treated as exogenous variables. Midwest/Northeast indicators are based on the U.S. Census Bureau's region definitions. House price growth rates are computed using the Federal Housing Finance Agency State House Price Indexes (seasonally adjusted, purchase-only index). Oil production is the U.S. Energy Information Administration's 2010 crude oil production, measured in barrels.

A.2 Additional Tables

Table A.3: Response of average employment growth (monthly average between April 2009 and March 2011) to state government losses and obligated ARRA money, fungibility not imposed, all coefficients reported

	Goods-producing	Services	Government
AID	0.15690	0.05978	0.06522
	(0.16930)	(0.04458)	(0.10440)
-LOSS	0.02801	0.01288	0.17357
	(0.10324)	(0.02619)	(0.06097)
Pre-recession level (in billions)	0.01906	0.00054	-0.00244
	(0.00784)	(0.00042)	(0.00404)
One-year lag employ. growth	0.09112	0.18034	-0.02912
	(0.09133)	(0.08220)	(0.14919)
Two-year lag employ. growth	-0.34888	-0.24193	-0.05327
	(0.15586)	(0.13949)	(0.16043)
Midwest/Northeast region	0.02510	-0.00290	0.00007
	(0.00636)	(0.00198)	(0.00356)
Oil-production per capita	0.17971	-0.10343	0.29637
	(0.42367)	(0.10742)	(0.23091)
House price grow (03-06)	-0.01776	-0.00160	0.00649
	(0.02984)	(0.00752)	(0.01600)
House price grow (06-08)	0.14498	0.01282	0.04359
	(0.03067)	(0.00860)	(0.01458)
Constant	-0.09104	-0.00285	0.01642
	(0.02912)	(0.00894)	(0.02018)
Reject fungibility (p-value)	0.481	0.619	0.604

Table A.4: Our benchmark estimates of jobs created/saved as a result of ARRA-spending component (in millions), with Congressional Budget Office and President's Council of Economic Advisors ranges shown for comparison

	Benchmark (this paper)		President Council of Economic Advisors		Congressional Budget Office	
	Low	High	Low	High	Low	High
2009:Q2			0.3	0.4	0.1	0.5
2009:Q3			1.0	1.1	0.3	1.2
2009:Q4			1.8	1.8	0.5	1.9
2010:Q1			2.2	2.6	0.6	2.7
2010:Q2			2.6	3.2	0.7	3.4
2010:Q3			2.7	3.5	0.7	3.6
2010:Q4			2.5	3.6	0.6	3.5
2011:Q1			2.4	3.6	0.6	3.3
Total employment (avg.)	0.1	1.5	1.9	2.5	0.5	2.5
Private sector (avg.)	-0.2	1.1				
Government (avg.)	0.2	0.6				

Notes: Benchmark low-high estimate is 90% confidence interval for national employment due to ARRA-spending component (averaged between April 2009 and March 2011). Other sources are Council of Economic Advisors (various quarterly reports) and Congressional Budget Office (various quarterly reports). The CEA and CBO estimates each include the effects of tax cuts.