Fiscal Spending Jobs Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act[†]

By Daniel J. Wilson*

This paper estimates the "jobs multiplier" of fiscal stimulus spending using the state-level allocations of federal stimulus funds from the American Recovery and Reinvestment Act (ARRA) of 2009. Because the level and timing of stimulus funds that a state receives was potentially endogenous, I exploit the fact that most of these funds were allocated according to exogenous formulary allocation factors such as the number of federal highway miles in a state or its youth share of population. Cross-state IV results indicate that ARRA spending in its first year yielded about eight jobs per million dollars spent, or $125,000 per job. (JEL E24, E62, H72, H75, R23)

> *Not for the first time, as an elected official, I envy economists. Economists have available to them, in an analytical approach, the counterfactual… They can contrast what happened to what would have happened. No one has ever gotten reelected where the bumper sticker said, 'It would have been worse without me.' You probably can get tenure with that. But you can't win office.*
>
> — US Representative Barney Frank, July 21, 2009[1]

T his paper analyzes the fiscal stimulus spending provided by the American Recovery and Reinvestment Act (ARRA) of 2009 and contrasts "what happened to what would have happened." It does so by exploiting the cross-sectional geographic variation in ARRA spending and the many exogenous factors that determined that variation. The use of cross-sectional variation, in contrast to most prior studies of the economic effects of fiscal policy which rely on time series variation, greatly mitigates the risk of confounding fiscal policy effects with effects from other macroeconomic factors, such as monetary policy, that are independent of the

geographic distribution of stimulus funds. In addition, because the level and timing of ARRA funds that a state receives is potentially endogenous with respect to its economic conditions, I make use of the fact that most of these funds were allocated according to statutory formulas based on exogenous factors such as the number of highway lane-miles in a state or the youth share of its population. I also utilize data on the initial announcements and obligations of ARRA funding by state, as opposed to actual outlays, to mitigate concerns about anticipation effects and implementation lags, the importance of which has been stressed in a number of recent studies.[2] Specifically, I provide instrumental variables (IV) estimates of the impact on employment of ARRA spending announcements, obligations, and outlays using instruments based on these formulary factors and controlling for variables that might be correlated with both the instruments and employment outcomes.

The ARRA was enacted into law in February 2009 amidst a great deal of economic and political debate. At roughly $800 billion, it was one of the largest fiscal stimulus programs in American history.[3] Proponents saw the stimulus package as a vital lifeline for an economy heading toward a second Great Depression. They pointed to projections from the White House and others suggesting that the stimulus package would create or save around 3.5 million jobs in its first two years. Critics claimed the massive cost of the ARRA would unduly swell the federal deficit while having minimal or even negative impact on employment and economic growth.

The policy debate over the effectiveness of the ARRA has centered around, and revived interest in, the long-standing economic debate over the size of fiscal multipliers. Ramey (2011a) surveys the literature on fiscal multipliers, pointing out that there is little consensus either theoretically or empirically on the size of the multiplier. As the quote at the beginning of the paper alludes to, the key challenge faced by researchers estimating the economic effects of fiscal policy is isolating changes in economic outcomes due solely to government spending from what would have occurred in the absence of that spending. This paper turns to cross-sectional geographic variation in government spending to identify fiscal effects, exploiting the fact that other potentially confounding nationwide factors such as monetary policy are independent of relative spending, and relative economic outcomes across regions. Other recent papers also have followed this approach. Nakamura and Steinsson (2010) use cross-region variation in US military spending to estimate an "open economy" fiscal multiplier, instrumenting for actual spending using a region's historical sensitivity to aggregate defense spending. Serrato and Wingender (2010) consider variation in federal spending directed to US counties, and take advantage of the natural experiment afforded by the fact that much federal spending is allocated based on population estimates that are exogenously "shocked" after each Decennial Census. Shoag (2010) estimates the multiplier associated with state-level government spending driven by exogenous shocks to state pension fund

---

[2] See, for example, Leeper, Walker, and Yang (2010), Ramey (2011b), and Mertens and Ravn (2010).

[3] When it was first passed, the ARRA was estimated to cost $787 billion over ten years. Most recent estimates put the cost at $821 billion, of which about two-thirds comes from increased federal government spending and one third from reduced tax revenues (see Congressional Budget Office 2011).

returns. Fishback and Kachanovskaya (2010) estimate a fiscal multiplier using variation across states in federal spending during the Great Depression. The results of Fishback and Kachanovskaya are particularly relevant here in that they, like this paper, investigate the fiscal multiplier during a time of considerable factor underutilization, when the multiplier should be at its largest according to traditional Keynesian theory. Fishback and Kachanovskaya find that government spending had a negligible impact on employment during the 1930s.

Fiscal multipliers estimated from cross-regional variation are, strictly speaking, "local" multipliers. That is, they correspond most closely to contexts in which output and factors of production are at least partially mobile across borders. Whether they are larger or smaller than national multipliers is not clear. To the extent that factor and goods mobility is greater among subnational regions than among countries, the local multiplier may be an upper bound on the national multiplier in non-tradable sectors—because factor mobility mitigates crowd-out of private sector production—but a lower bound in tradable goods sectors, as the benefits of the local demand shock spillover to other regions (see Moretti 2010).[4] Also, the multipliers estimated from cross-sectional studies may be larger than a national multiplier because of the independence between the geographic allocation of federal spending and the geographic allocation of the financing of that spending. For instance, suppose a single region received 100 percent of federal government spending. The burden imposed by that spending on the federal government's budget constraint will be shared by taxpayers in all regions. In this sense, cross-sectional studies provide estimates of the multiplier associated with "windfall" government spending, which could have a higher or lower short-run multiplier than that of deficit-financed spending (Clemons and Miran 2011). In a standard Neoclassical model, for instance, deficit-financed spending leads households to increase labor supply and hence output as forward-looking households recognize that increased government spending necessitates increased future taxes. In that model, GDP multipliers based on windfall spending, such as those estimated in this paper, would be smaller than multipliers based on deficit-financed spending, such as the national multiplier.

Of course, the fact that the local multiplier may not equal the national multiplier does not mean that the local multiplier is not of independent interest, nor does it mean that the local multiplier cannot inform the debate surrounding the effectiveness of federal stimulus. In the United States and many other countries with federalist systems, a large share of federal spending comes in the form of regional transfers. The economic impact of these transfers is of first-order importance. In addition, this paper provides evidence on how the employment effects of ARRA spending changed over time. The factors potentially causing a gap between the local and national multiplier (interregional factor mobility and the extent to which agents are forward-looking) are likely to be fairly constant over time, implying that the national effects evolved over time similarly to the local effects.

---

[4] Ilzetzki, Mendoza, and Végh (2010), in their cross-country panel study, find evidence that the fiscal spending multiplier is lower in open economies than in closed economies. To the extent that subnational regions within the United States are more open than the national economy, this result suggests that the local multiplier estimated for these regions may indeed be a lower bound for the national multiplier.

Since the ARRA's passage, a number of studies have sought to measure its economic effects. The methodologies used in these studies can be divided into two broad categories. The first methodology employs a large-scale macroeconometric model to obtain a baseline, no-stimulus forecast and compares that to a simulated forecast where federal government spending includes the ARRA. This is the methodology used in widely cited reports by the Congressional Budget Office (CBO) (see, e.g., CBO 2011), the White House's Council of Economic Advisers (CEA) (e.g., CEA 2011), private forecasters such as Macroeconomic Advisers, IHS Global Insight, and Moody's Economy.com, as well as a number of academic studies.[5] The key distinction between that methodology and the one followed in this paper is that the former does not use observed data on economic outcomes following the start of the stimulus. Rather, it relies on a macroeconometric model, the parameters of which, including its fiscal spending multiplier(s), are estimated using historical data prior to the ARRA (or pulled from the literature which estimated them using historical data).[6]

The second methodology is an attempt to count the jobs created or saved by requiring "prime" (or "first-round") recipients of certain types of ARRA funds to report the number of jobs they were able to add or retain as a direct result of projects funded by the ARRA. These counts are aggregated up across all reporting recipients by the Recovery Accountability and Transparency Board (RATB)—the entity established by the ARRA and charged with ensuring transparency with regard to the use of ARRA funds—and reported online at www.recovery.gov and in occasional reports to Congress.[7] The number of jobs created or saved, and any fiscal multiplier implied by such a number, reflects only "first-round" jobs tied to ARRA spending, such as hiring by contractors and their immediate subcontractors working on ARRA funded projects, and excludes both "second-round" jobs created by lower-level subcontractors and jobs created indirectly due to spillovers such as consumer spending made possible by the wages associated with these jobs and possible productivity growth made possible by ARRA-financed infrastructure improvements. By contrast, the methodology of this paper uses employment totals as reported by the Bureau of Labor Statistics, and therefore all direct and indirect jobs created by the ARRA should be reflected in the results. Furthermore, only 55 percent of ARRA spending is covered by these recipient reporting requirements (see CEA 2010, 27).

The methodology I employ in this paper is distinct from the above two methodologies in that it uses both observed data on macroeconomic outcomes—namely, employment—and observed data on actual ARRA stimulus spending. This paper was the first, to my knowledge, to exploit the cross-sectional, geographic variation in ARRA spending to estimate its economic effects. However, a number of other studies also have followed a similar approach. One is the paper by Chodorow-Reich et al. (forthcoming), which investigates the employment effects of the ARRA's Medicaid spending, finding that such spending generated 38 job-years per million dollars, or about $26,000 per job. This paper and Chodorow-Reich et al. (forthcoming) share

---

[5] See, for example, Cogan et al. (2009), Blinder and Zandi (2010), and Drautzburg and Uhlig (2010).

[6] CEA (2010) also estimates the ARRA's economic impact using a VAR approach that compares forecasted post-ARRA outcomes (employment or GDP), based on data through 2009:I, to actual post-ARRA outcomes.

[7] For more details and discussion of these data on ARRA job counts, see Government Accountability Office (2009) and CBO (2011).

in common the use of states' pre-ARRA Medicaid expenditures as an instrument for ARRA spending done by the Health and Human Services department. (Although, as I show later in the paper, my baseline empirical results are qualitatively unchanged if I exclude this instrument.) However, this paper uses that instrument along with instruments for other departments' ARRA spending in order to estimate the overall employment effects of ARRA spending, while Chodorow-Reich et al. (forthcoming) focus on the narrower question of the specific impact of the ARRA's Medicaid fiscal relief fund. Furthermore, Chodorow-Reich et al. (forthcoming) rely on ARRA payments (outlays) to measure stimulus spending whereas this paper utilizes data on announcements and obligations of ARRA funds, which I argue below are likely to better reflect the funding amounts that agents anticipate. A second related paper is Feyrer and Sacerdote (2011), which utilizes state and county level variation in ARRA outlays and employment outcomes. Another is Conley and Dupor (2011), which uses cross-state variation in ARRA payments and obligations, instrumenting with ARRA highway funding and states' reliance on sales taxes, to estimate ARRA's spending overall employment effect as well as its effect in selected subsectors. A discussion of how the empirical approaches and results from these latter two studies compare to this paper is provided in Section V.

The remainder of the paper is organized as follows. The next section discusses the empirical methodology and describes the data used in the analysis. The baseline empirical results, showing the ARRA's impact on employment as of its one-year mark, are presented and discussed in Section III. Section IV considers how the employment effect varied across sectors and over time. In Section V, I discuss the implications of these results and compare them with other studies relating to the ARRA and fiscal stimulus in general. Section VI offers some concluding remarks.

I. Empirical Model and Data

A. *Baseline Empirical Model*

I estimate the following cross-state instrumental variables regression:

$$(1a) \qquad (L_{i,T} - L_{i,0}) = \alpha + \beta S_{i,T} + \mathbf{X}_{i,0}\Gamma + \varepsilon_{i,T}$$

$$(1b) \qquad S_{i,T} = \delta + \lambda(L_{i,T} - L_{i,0}) + \mathbf{X}_{i,0}\Theta + \mathbf{Z}_{i,0}\Phi + \upsilon_{i,T}.$$

$(L_{i,T} - L_{i,0})$ is the change in employment, scaled by 2009 population, from the initial period when the stimulus act was passed ($t = 0$) to some later period ($t = T$). $S_{i,T}$ is cumulative ARRA spending per capita in state i as of period T. $\mathbf{X}_{i,0}$ is a vector of control variables ("included" instruments). $\mathbf{Z}_{i,0}$ is a vector of ("excluded") instruments.

I will refer to β as the fiscal jobs multiplier. Formally, β represents the marginal effect of per capita stimulus spending on employment change from period 0 to T:

$$(2) \qquad \beta \equiv \frac{\partial((L_{i,T} - L_{i,0})/POP_{i,0})}{\partial(S_{i,T}^{\$}/POP_{i,0})} = \frac{\partial(L_{i,T} - L_{i,0})}{\partial S_{i,T}^{\$}},$$

where $POP_{i,0}$ is state population in 2009, and $S_{i,T}^\$$ is the level of cumulative stimulus spending in the state ($S_{i,T}^\$ \equiv S_{i,T} \times POP_{i,0}$). The reciprocal of β represents the stimulus cost per job created or saved.

B. *Dependent Variable and Controls*

Except when otherwise noted, monthly state employment is measured using the Bureau of Labor Statistics' (BLS) Current Employment Statistics (CES) payroll survey data on seasonally-adjusted employment in the total nonfarm sector.[8]

I include five control variables in each regression. These control variables are included because they are likely to be both good predictors of economic outcomes over the post-ARRA period, and could be correlated with the instruments used to predict ARRA spending. Following Blanchard and Katz's (1992) empirical model of state employment growth, I control for lagged employment growth and the initial level of employment. Specifically, I include the change in employment per capita from the start of the recession (December 2007) to when the ARRA was enacted (February 2009), and the initial level of employment per capita in February 2009.

One unique feature of the 2007–2009 recession was the rapid run-up in house prices preceding the recession followed by a sharp correction. The magnitude of the run-up and subsequent correction varied considerably across states and one might be concerned that the house price run-up could be correlated with one or more of the formula factors used as instruments. For this reason, I control for the percentage change in the state's house prices from 2003 to 2007 using the state-level House Price Index from the Federal Housing Finance Agency (FHFA).

The last two controls account for specific features of the ARRA. The first is the change, from 2005 to 2006, in a three-year trailing average of personal income per capita. This variable is included because it directly enters the formula determining the state allocations of ARRA "Fiscal Relief" funds, which come from the Department of Health and Human Services (HHS) and were meant to help states pay Medicaid expenses.[9] The last control is estimated ARRA tax benefits received by state residents. This variable is the sum of estimated tax benefits from the ARRA's "Making Work Pay" (MWP) payroll tax cut and the ARRA's increase of the income thresholds at which the Alternative Minimum Tax (AMT) becomes binding. Following the Center for Budget and Policy Priorities (CBPP), the MWP benefits are estimated

---

[8] Preliminary CES data, which come out approximately two months after the month in question, are based on a payroll survey of about 400,000 business establishments and some model-based adjustments for establishment entry and exit. These data are revised annually to incorporate information on state employment levels from state UI records and updated seasonal adjustment factors. As of the time of this writing, the last benchmark revision was done in March 2011, revising state employment counts for months from October 2009 through September 2010. Thus, the employment data used in this paper through September 2010 incorporate the benchmark revisions and are unlikely to be substantially revised in the future. However, employment data for October 2010 through March 2011 are preliminary and so the results presented herein for these months should be considered tentative.

[9] The hold-harmless component of the ARRA's Medicaid funds calls for states whose FY2009 Federal Medicaid Assistance Percentage or FMAP (an inverse function of mean personal income per capita from 2004–2006) is greater than FY2008 FMAP (an inverse function of personal income per capita from 2003–2005) to receive Medicaid funds based on FY2008 FMAP (plus other adjustments to this percentage specified in the ARRA). So the hold-harmless component of a state's ARRA Medicaid funds is increasing in FY2009 FMAP—FY2008 FMAP, which in turn is a function of the change in the three-year moving average of personal income per capita lagged three years.

by taking each state's share of the national number of wage/salary earners making less than \$100,000 for single filers and less than \$200,000 for joint filers (roughly the levels above which the MWP benefit phases out), as of 2006, and multiplying by the total cost of MWP tax cuts (\$116.2 billion over ten years, according to CEA 2010). Similarly, using state-level data from the Tax Policy Center on each state's share of national AMT payments, as of 2007, I estimate AMT benefits by multiplying that share by the total cost of the AMT adjustment (\$69.8 billion, according to CEA 2010).

I also report results below using state population as an additional control variable. This variable turns out to be statistically insignificant and has virtually no effect on the estimated stimulus impact and hence is not included in the baseline specification.

C. *Stimulus Spending Measures*

The data on ARRA spending by state come from the Recovery.gov website. The website reports three different spending measures: (i) funding announcements, (ii) funding obligations, and (iii) final payments (outlays).[10] Announcements are reported in periodic Funding Notification Reports, while obligations and payments are reported in weekly Financial and Activity Reports. Recovery.gov provides both current and past Financial and Activity Reports, allowing one to construct cumulative-to-date measures of obligations and payments for each month from April 2009 onward. Unfortunately, the website does not provide archived Funding Notification Reports, making it impossible to compile data on announcements over time from the information currently provided on the website. However, for this paper I began routinely downloading the Funding Notification Reports starting in August 2009 and hence am able to measure ARRA funding announcements from August 2009 onward.

For all of the analyses in this paper, I exclude spending done by the Department of Labor (DOL), which primarily is funds sent to state governments to pay for extended and expanded unemployment insurance (UI) benefits, for two reasons. First, announcements of DOL funding are not reported by state on Recovery.gov. Second, and most importantly, this type of spending in a given state is driven almost entirely by the change in the state's unemployment rate. There is virtually no source of exogenous variation to use as an instrument for this variable. The numbers reported in the remainder of the paper reflect non-DOL ARRA spending only. (DOL spending accounted for 14 percent (\$66.5 billion) of total obligations through March 2011.)

How do these three measures of spending differ conceptually? Figure 1 provides a schematic that depicts how these three measures are related in terms of the timing of spending.

The ARRA provided additional budget authority to federal agencies to obligate funds above the levels provided in the previously enacted fiscal year 2009 budget. The legislation also specified formulas or mechanisms for how agencies should

---

[10]Recovery.gov provides both recipient-reported data and agency-reported data. Because the recipient-reported data only cover a little over half of all ARRA spending, I use the agency-reported data, which covers all ARRA spending. It should be noted that for each measure of spending, not all funds are reported separately by state. As of the end of 2010, 18 percent of announcements, 12 percent of obligations, and 12 percent of payments were not separated by state. For the remainder of the paper, I will use and discuss only the state-allocated data.

```
┌──────────────────────────────────────────────────────────────┐
│      Funds authorized to a given federal agency by ARRA legislation      │
└──────────────────────────────────────────────────────────────┘
        ┌───────────────────────────┬───────────────────────────┐
┌───────────────────────────────┐   ┌───────────────────────────────┐
│  Funds announced as available to applicants,  │   │       Unannounced funds       │
│    conditional on satisfying requirements     │   │                               │
└───────────────────────────────┘   └───────────────────────────────┘
        └───────────────────────────┬───────────────────────────┘
              ┌─────────────────────────────────────┐
              │   Funds obligated to specific recipients   │
              └─────────────────────────────────────┘
                      ┌───────────────────────────────┐
                      │    Funds paid out to recipients    │
                      └───────────────────────────────┘
```
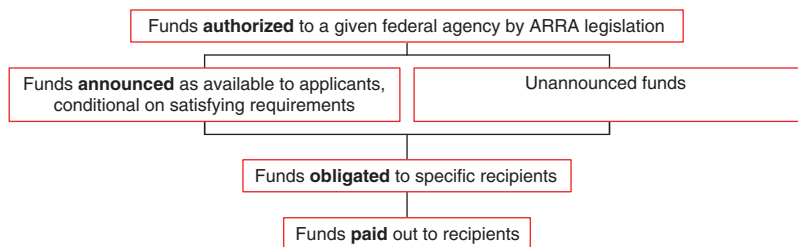
Figure 1. Flow of ARRA Spending

allocate those funds to recipients. Many of these formulas are simply the pre-existing formulas used by the agency to allocate non-ARRA funds; in essence, the ARRA just stepped up the funding amounts without altering the allocation mechanism. Based on that authorization, agencies subsequently *announce* how much funding is available for eligible recipients in each state, though a small portion of authorized funds are never announced.[11] Whether they are announced or not, authorized funds are eventually awarded, or "*obligated*," to recipients. For example, the Department of Transportation (DOT) might award a contract to a construction firm or municipal agency at which point the DOT is said to have obligated those funds to that recipient. Finally, when recipients satisfy the terms of their contracts, the agency actually pays out the funds.

Based on CBO estimates, by the end of 2010 nearly 90 percent of the expected 10-year ARRA spending total had been obligated and 65 percent had been paid out. The progression of spending can be seen in Figure 2, which shows state-allocated ARRA funding announcements, obligations, and payments from April 2009 through March 2011.[12] By the end of this period, announcements, obligations, and payments were $282.9 billion, $337.0 billion, and $229.3 billion, respectively.

The ARRA spending (excluding DOL spending) is spread over dozens of separate federal agencies, though three agencies in particular account for the bulk of it. The disaggregation across major agencies is shown in Table 1. Through March 2011, the Departments of Education (ED), Health and Human Services (HHS), and Transportation (DOT) were responsible for 64 percent of the spending announcements, 73 percent of obligations, and 79 percent of payments. For this reason, while I include all non-DOL stimulus spending in the analysis, my instrumental variables approach to identifying the exogenous component of stimulus spending will focus on the formula factors that go into the allocations of ED, HHS, and DOT spending.

Although I report regression results for all three measures of spending, it is worth discussing the relative merits of each as a measure of fiscal stimulus. The advantage

[11] The glossary entry for "Funds Announced by an Agency" on recovery.gov states: "Not all available funds are announced publicly. For example, the funds going to a project started prior to the Recovery Act that are commingled with the project's Recovery funds will not be announced publicly before being made available to a recipient."

[12] Note that total announcements are observed only for August 2009 onward. Recovery.gov does not provide archived Funding Notifications (the source of announcements data) and August 2009 was the first month in which I began regularly downloading the Funding Notification reports. For some agencies, however, announcements are known for earlier months because their August 2009 Funding Notifications indicated that the reported level of announced funds is "as of" a specified earlier month. The earlier "as of" month is reflected in the announcements-by-agency levels shown in Figure 4.
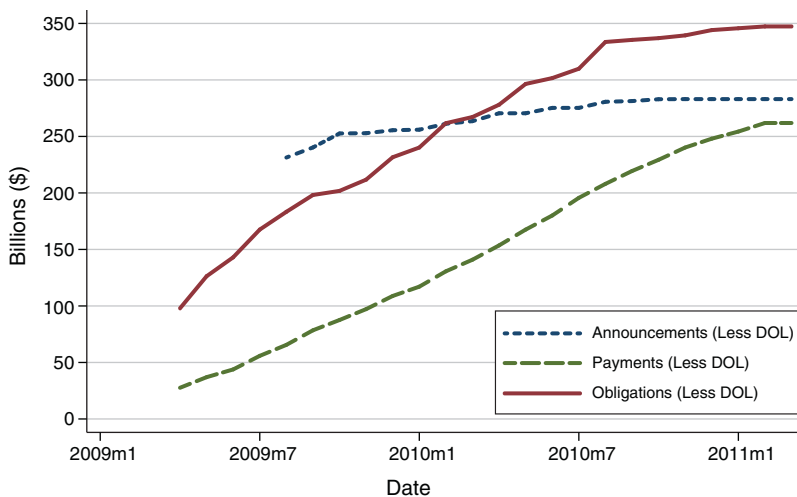
FIGURE 2. ARRA SPENDING MEASURES OVER TIME

TABLE 1—AGENCY TOTALS (BILLIONS) AND PERCENTAGES
MARCH 2011

	Announcements		Obligations		Payments	
Dept. of Education (ED)	89.1	(31.5)	94.6	(27.2)	74.2	(28.4)
Dept. of Transportation (DOT)	34.7	(12.3)	42.7	(12.3)	25.2	(9.6)
Other	103.6	(36.6)	96.8	(27.9)	61.3	(23.4)
Dept. of Health and Human Services (HHS)	55.6	(19.6)	113.1	(32.6)	101.1	(38.6)
Total (excluding Dept. of Labor)	283.0	(100.0)	347.3	(100.0)	261.8	(100.0)

of announcements and obligations relative to payments is that the former two mea-sures are likely to lead (affect) employment and other economic activity, whereas payments are likely to lag activity. For instance, private contractors are most likely to make job hiring or retention decisions when they begin a project, which will occur after they have been awarded a contract. If the contract is awarded directly by a federal agency, the timing of the contract award will be reflected in the timing of the obligations data. If the contract is awarded by a state or local government agency, which received funding from the ARRA, the contract award will occur at some point after the announcement and obligation of those funds to the state or local agency. A payment will not occur until the contract is completed, and possibly even later if there are bureaucratic delays in disbursements.[13]

Announcements generally lead obligations by several months. For job creation/retention of private contractors funded directly by federal agencies, obligations are likely the most relevant measure because they reflect contract awards to a specific contractor. For job creation/retention decisions by state and local governments or

---

[13] See Leduc and Wilson (forthcoming) for discussion and evidence on the long implementation lags between funding announcements and payments for government infrastructure spending.

decisions by contractors funded by state or local government agencies, announcements may be the most relevant measure since the timing of announcements reflect when a state or local government first learns that it is eligible to receive the funds, and it can then act upon that information in its budgeting and personnel decisions.[14] Thus, in terms of obligations versus announcements, obligations have the advantage of reflecting only guaranteed funding to specified recipients (as opposed to notification of funding eligibility) and, at least for private projects funded directly by federal agencies, may be timed closer to the start of project when hiring is most likely to occur. From a more general theoretical perspective, however, announcements have the advantage of more closely reflecting the timing of agents' expectations about future stimulus spending in their state. Recent papers by Ramey (2011b), Leeper, Walker, and Yang (2010), and Mertens and Ravn (2010) have demonstrated the empirical importance of measuring the timing of initial announcements and/or anticipation of government spending shocks. For this reason, I treat announcements as the preferred measure of stimulus spending throughout the paper.

D. Instruments

As mentioned earlier, the stimulus variable, $S_{i,T}$, may well be endogenous ($\lambda \neq 0$). There are two potential sources of endogeneity. First, some of the components of $S_{i,T}$ are explicitly functions of current economic conditions. For example, consider the formula determining the state allocation of spending from HHS' "Fiscal Relief Fund," which is meant to help state governments pay for Medicaid expenses. Each state's per capita receipts from this Fund depend on three factors: (1) the current federal Medicaid share (which is a function of pre-ARRA income per capita), (2) the "hold-harmless" component (a function of 2006–2007 growth in a three-year moving average of state income per capita, which I control for in the regressions), and (3) the change in the unemployment rate from the beginning of the recession through February 2009. These factors determining ARRA Fiscal Relief funds may also be correlated with post-stimulus economic conditions—for example, states with a rapid pre-stimulus increase in the unemployment rate may be likely to rebound more quickly than other states because the rapid increase might suggest those states tend to enter and exit recessions earlier than others. However, note that if these factors are controlled for directly in $\mathbf{X}_i$, then this source of endogeneity should be eliminated.

A second potential source of endogeneity, especially for obligations and payments, is that the level and timing of ARRA spending going to any given state is partly a function of how successful the state government is at soliciting funds from federal agencies. Most of the state allocation of funding announcements is exogenously determined by formulas, but much of obligations and payments are allocated at the discretion of the federal agencies as they review whether states have satisfied so-called "maintenance of effort" (MOE) requirements and what the states' plans are for how they intend to spend the money. States with unfavorable MOEs or

---

[14] Note that state and local governments generally are able to avoid any temporary cash flow shortage through short-term borrowing (e.g., issuing revenue anticipation notes or warrants).

spending plans may receive funding later or not at all (e.g., DOT funds have a "use it or lose it" requirement).[15] States that are more successful in soliciting funds and starting projects may also be better-run state governments, and better-run states may be more likely to have positive outcomes regardless of the stimulus funds. One can address this source of endogeneity via instrumental variables.

I instrument for actual ARRA spending (measured by announcements, obligations, or payments) by state, $S_{i,T}$, using instruments based on the formulary allocation factors. The vast majority of ARRA spending was allocated to states according to statutory formulas, some of which were unique to the ARRA and some of which were pre-existing formulas already used for non-ARRA funds. In particular, I construct three instruments based on the formulas used by the three federal agencies responsible for the bulk of ARRA spending: the Departments of Transportation (DOT), Education (ED), and HHS.[16]

Department of Transportation Instrument.—Around three-quarters of the DOT's $40 billion of ARRA funding was allotted to the Federal Highway Administration (FHWA), which funds construction and repair of highways and bridges. The ARRA specified that 50 percent of these FHWA funds be allocated to states based on the pre-existing allocation formula used by the FHWA's Surface Transportation Program (STP). With some exceptions, this formula is simply a weighted average of three factors, each measured with a three-year lag: the state's lane-miles of federal-aid highways, its estimated vehicle miles traveled on federal-aid highways, and its estimated payments (primarily from a motor fuels excise tax) into the federal highway trust fund. Thus, this portion of the DOT's ARRA funds in 2009 were allocated to states roughly in proportion to the values of these three factors in 2006. The other 50 percent was set to be allocated in exactly the same proportion as FHWA used for non-ARRA funds (obligation limitations) in 2008, of which the STP program is a major component. Thus, those STP program formula factors help explain a large share of the FHWA's, and hence DOT's, ARRA spending by state. Both the three-year lag and the likely weak association between such factors and a state's current economic conditions suggest an instrument constructed from these factors should satisfy the exclusion restriction of the IV estimator. I construct such an instrument by taking the fitted value from regressing the DOT's ARRA obligations in 2009 on total lane miles of federal-aid highways in 2006, total vehicle miles traveled on federal-aid highways in 2006, estimated tax payments attributable to highway users paid into the federal highway trust fund in 2006, and FHWA obligation limitations in 2008.

Department of Education Instrument.—Over half of the ED's ARRA spending comes from the $54 billion State Fiscal Stabilization Fund (SFSF) set up by the legislation. The ARRA specified that 61 percent of the fund would be allocated to

states in proportion to their school-aged (5–24 years old) population and the other 39 percent would be allocated in proportion to total population. This implies that the per capita stimulus that any given state can be expected to receive over the life of the ARRA is a simple linear function of the state's school-aged population share, making this share a natural instrument for actual per capita ARRA education spending received by a state.[17] The youth population share will satisfy the exclusion restriction as long as states' post-ARRA employment growth, conditional on pre-ARRA employment trends and the other controls, is not systematically related to states' youth population share.

Health and Human Services Instrument.—The ARRA also set up an $87 billion State Fiscal Relief Fund meant to help states pay their Medicaid expenses. The formula for allocating these funds to states was described above. Among other components, the ARRA temporarily increased the state-specific percentages of Medicaid expenditures reimbursed by HHS by 6.2 percentage points for all states. This means that states that, prior to the ARRA, had higher Medicaid expenditures per capita (either because of having more low-income families per capita or because they have more generous Medicaid programs) will tend to receive more ARRA HHS funds per capita. Therefore, pre-ARRA state Medicaid expenditures should be a good predictor of ARRA HHS funds. And because pre-ARRA state Medicaid expenditures are based on conditions (generosity of the state's Medicaid program and income per capita from earlier years) prior to the ARRA, they should be orthogonal to post-ARRA economic conditions, making them suitable instruments for realized ARRA health spending. Chodorow-Reich et al. (forthcoming) use this instrument as well in their study of the economic impact of the State Fiscal Relief Fund.

II. Baseline Results

A. *A Simple Illustrative Comparison of High and Low Stimulus States*

Before estimating the baseline empirical model described above, it is useful to consider the results of a simple exercise that illustrates the key sources of variation used to identify that baseline model.[18] First, I separate states into quintiles based on predicted ARRA spending—that is, the fitted values from regressing ARRA spending on the three instruments described above. I then compare employment over time for the top quintile to that of the bottom quintile. Table 2 shows the full ranking of states by predicted ARRA announcements. The ranking is similar if done according to obligations or payments. The table also shows each state's values for actual announcements and each of the three instruments.

---

[17] The SFSF funding allocated to state i, S_i, is determined according to the formula, $S_i = \bar{S}\left[0.61\left(\frac{y_i}{\bar{y}}\right) + 0.39\left(\frac{p_i}{\bar{p}}\right)\right]$, where y_i denotes school-aged population, p_i denotes total population, and bars above variables indicate national totals. A state's per capita SFSF authorized funding is then, $\frac{S_i}{p_i} = 0.39\left(\frac{\bar{S}}{\bar{p}}\right) + 0.61\left(\frac{\bar{S}}{\bar{y}}\right)\left(\frac{y_i}{p_i}\right)$.

[18] A similar exercise was done in Mian and Sufi's (2010) study of the effect of the Cash for Clunkers program on auto purchases.

TABLE 2—SELECTED VARIABLES, AS OF FEBRUARY 2010
STATES RANKED BY PREDICTED ARRA ANNOUNCEMENTS

Rank	State	Predicted announcements	Actual announcements	HHS instrument	ED instrument	DOT instrument
1	WY	1,542	1,096	48	0.273	296
2	ND	1,461	1,331	48	0.291	270
3	AK	1,409	2,168	85	0.296	224
4	SD	1,264	1,290	47	0.278	229
5	MT	1,231	1,476	46	0.269	225
6	VT	1,196	1,026	105	0.264	165
7	NM	957	1,187	81	0.274	127
8	RI	937	828	99	0.265	109
9	MS	903	846	68	0.29	120
10	NE	876	747	52	0.283	131
11	ME	868	1,046	93	0.246	104
12	AR	865	746	66	0.267	122
13	NY	864	803	138	0.262	57
14	WV	858	883	73	0.246	120
15	OK	847	755	55	0.276	124
16	DE	821	899	69	0.258	111
17	KS	818	634	47	0.284	122
18	ID	815	996	43	0.291	122
19	MO	811	701	68	0.27	106
20	IA	806	756	51	0.275	118
21	LA	799	823	72	0.282	95
22	MN	788	628	75	0.27	94
23	AL	773	653	54	0.27	109
24	CT	760	571	75	0.266	89
25	KY	755	651	65	0.263	98
26	TN	740	773	70	0.259	91
27	MA	737	696	96	0.258	67
28	IN	735	656	49	0.277	102
29	SC	735	925	55	0.264	101
30	PA	731	604	78	0.258	82
31	WI	708	554	53	0.272	94
32	OH	706	595	69	0.267	81
33	NH	700	1,028	54	0.262	94
34	TX	698	723	51	0.286	89
35	MI	697	541	57	0.278	85
36	GA	682	541	43	0.28	94
37	AZ	677	769	62	0.27	79
38	NC	669	505	62	0.267	78
39	IL	649	719	60	0.277	72
40	NJ	647	347	63	0.257	75
41	MD	645	805	59	0.265	76
42	CA	633	564	59	0.279	69
43	OR	631	693	46	0.256	87
44	UT	631	659	30	0.326	78
45	VA	618	641	39	0.266	88
46	HI	613	1,057	52	0.251	79
47	WA	605	968	53	0.259	74
48	CO	573	680	36	0.264	80
49	FL	547	553	45	0.244	72
50	NV	539	433	29	0.264	78

The results of the exercise are shown in Figure 3, which compares median employment across the ten states in the top quintile of predicted stimulus to that of the ten states in the bottom quintile. Each state's employment is scaled by its average 2008 value to put all states on similar scales. Employment was falling sharply for both sets of states prior to the ARRA's enactment in February 2009 (indicated by the vertical

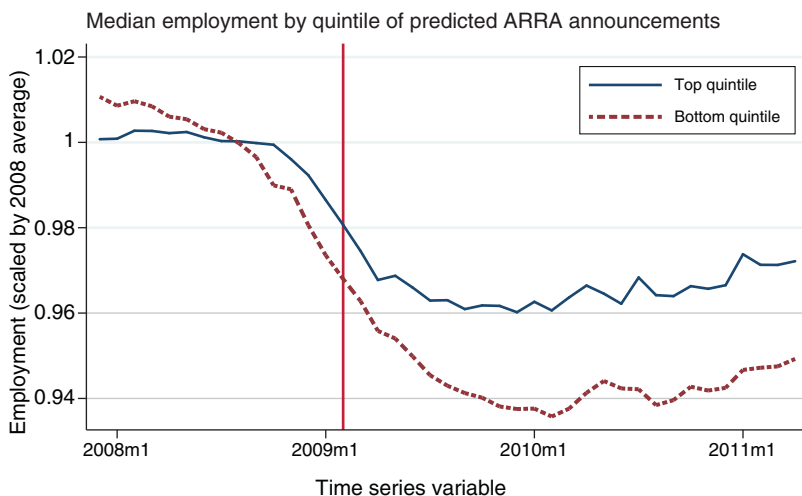Median employment by quintile of predicted ARRA announcements

FIGURE 3. EMPLOYMENT IN HIGH VERSUS LOW PREDICTED STIMULUS STATES

line), though it had been falling for longer for the bottom quintile. Employment began to stabilize for the high predicted-stimulus states—that is, those whose pre-ARRA formula factors (instruments) made them more likely to receive high levels of ARRA spending—around mid-2009, while employment for the low predicted-stimulus states continued to fall until early 2010. As a result, what was a small gap between the high and low stimulus states when the ARRA was enacted widened substantially in subsequent months. In particular, a large gap is evident by February 2010, one year after the Act was passed.

This simple comparison illustrates the basic idea of the more comprehensive empirical model estimated below, namely, comparing employment outcomes across states differentiated by how much ARRA spending they were predicted to receive based on pre-ARRA formula factors. However, the full empirical model will do this across all 50 states and will control for a number of other factors, such as the different trends in employment leading up to the ARRA as is evident in Figure 3.

B. *Estimates of the Baseline Model*

Summary statistics for the dependent variables, independent variables, and instruments used in the analysis are shown in Table 3. Table 4 provides the first-stage results for the IV estimation.[19] Column 1 shows the results when announcements are used as the measure of stimulus; columns 2 and 3 give the results when obligations and payments, respectively, are used. Each of the three instruments, which are based on pre-ARRA factors, is found to be positively related to the subsequent allocation of ARRA spending. The DOT and HHS instruments are statistically significant for predicting all three measures of spending. The ED instrument, youth population share, is statistically

---

[19] The stimulus variables are expressed here in thousands instead of millions to ease display of the coefficient values.

TABLE 3—SUMMARY STATISTICS, SAMPLE PERIOD: FEBRUARY 2009–FEBRUARY 2010

	Mean	SD	Min.	Max.	N
Panel A. Dependent variables					
Change in employment (p.c.), total nonfarm	−0.0151	0.0051	−0.0276	−0.0019	50
Change in employment (p.c.), private employment	−0.0145	0.0046	−0.0272	−0.0042	50
Change in employment (p.c.), S&L government	−0.0008	0.0010	−0.0028	0.0018	45
Change in employment (p.c.), construction	−0.0030	0.0019	−0.0128	0.0006	44
Change in employment (p.c.), manufacturing	−0.0034	0.0014	−0.0073	−0.0008	47
Change in employment (p.c.), educ. and health	0.0005	0.0006	−0.0008	0.0020	49
Change in the unemployment rate	0.0143	0.0077	0.0000	0.0430	50
Panel B. Explanatory variables					
Announcements (p.c.)	811.4	303.8	347.0	2,167.8	50
Obligations (p.c.)	903.9	199.1	636.8	1,802.8	50
Payments (p.c.)	431.8	85.7	269.3	644.1	50
Dec07–Feb09 employment (p.c.) trend, total nonfarm	−0.0223	0.0106	−0.0550	−0.0005	50
Dec07–Feb09 employment (p.c.) trend, S&L government	−0.0003	0.0008	−0.0025	0.0015	45
Dec07–Feb09 employment (p.c.) trend, total private	−0.0220	0.0101	−0.0536	−0.0004	50
Dec07–Feb09 employment (p.c.) trend, construction	−0.0039	0.0030	−0.0133	0.0001	47
Dec07–Feb09 employment (p.c.) trend, manufacturing	−0.0051	0.0025	−0.0135	−0.0016	47
Dec07–Feb09 employment (p.c.) trend, educ. and health	0.0008	0.0008	−0.0011	0.0025	49
Dec07–Feb09 employment (p.c.) trend, unemployment	0.0314	0.0113	0.0100	0.0540	50
Feb09 employment (p.c.) level, total nonfarm	0.4481	0.0415	0.3773	0.5676	50
Feb09 employment (p.c.) level, S&L government	0.0706	0.0126	0.0495	0.1166	45
Feb09 employment (p.c.) level, total private	0.3682	0.0362	0.2928	0.4490	50
Feb09 employment (p.c.) level, construction	0.0224	0.0057	0.0139	0.0459	47
Feb09 employment (p.c.) level, manufacturing	0.0419	0.0155	0.0157	0.0813	47
Feb09 employment (p.c.) level, educ. and health	0.0649	0.0150	0.0366	0.0983	49
Feb09 employment (p.c.) level, unemployment rate	0.0759	0.0182	0.0410	0.1200	50
Change in PI 3-yr moving average (p.c.), 2005 to 2006	0.0008	0.0005	−0.0002	0.0025	50
Tax benefits (p.c.)	563.7	110.5	434.6	921.0	50
2003–2007 house price growth	0.4	0.2	0.0	0.9	50
Panel C. Instruments					
DOT instrument	111.0	51.497	57.326	296	50.0
School-aged population share	0.270	0.014	0.244	0	50.000
FY 2007 p.c. Medicaid spending × 6.2 percent	61.9	20.1	29.2	138	50.0

significant for predicting payments but not for announcements or obligations, which is likely due to the positive collinearity between ED and the other two instruments.

The second-stage IV results along with OLS results for the baseline model — equations (1a) and (1b) — are shown in Table 5. Each column represents a separate regression. The dependent variable in each regression is the change in total nonfarm employment per capita (using 2009 population). The initial period ($t = 0$) for these regressions is February 2009, the month in which the ARRA was enacted. For the purposes of these baseline results, I choose the end period ($t = T$) to be February 2010. This choice is basically arbitrary—below I present the fiscal jobs multiplier estimates for other end-months—though February 2010 is of particular interest given that many studies of the short-run fiscal multiplier focus on the multiplier one year after the initial government spending shock. Bold coefficients are statistically significant at the 10 percent level or below. In addition to the ARRA spending variables, the explanatory variables include the 2005 to 2006 change in a three-year average of real personal income per capita (a factor in the allocation of HHS funds), an estimate of the per capita ARRA tax benefits going to the state, the change in employment from December 2007 to February 2009 (a measure of the

Table 4—First-Stage Results, as of February 2010

	Announcements (thous. per cap) β/SE	Obligations (thous. per cap) β/SE	Payments (thous. per cap) β/SE
HHS instrument (thous. per cap)	**3.838**	**6.683**	**3.221**
	(1.672)	**(1.060)**	**(0.594)**
DOT instrument (thous. per cap)	**3.402**	**2.359**	**0.542**
	(0.758)	**(0.480)**	**(0.269)**
ED instrument	2.391	1.473	**1.338**
	(2.031)	(1.287)	**(0.721)**
Change in PI moving average	**−128.698**	**−91.903**	**−63.400**
	(68.394)	**(43.347)**	**(24.283)**
Tax benefits (thous. per cap)	**−0.602**	0.017	0.124
	(0.335)	(0.212)	(0.119)
Dec07–Feb09 trend	**7.416**	−0.451	−1.948
	(3.639)	(2.306)	(1.292)
Feb09 level	−0.340	−0.043	0.186
	(0.898)	(0.569)	(0.319)
2003–2007 house price growth	**0.729**	**0.288**	**0.106**
	(0.177)	**(0.112)**	**(0.063)**
Constant	0.025	−0.209	−0.376
	(0.735)	(0.466)	(0.261)
Observations	50	50	50
R^2	0.714	0.732	0.547

Table 5—Second-Stage Results, as of February 2010

Dependent variable: Change in employment:population ratio, February 2009–February 2010

	OLS β/SE	IV β/SE	OLS β/SE	IV β/SE	OLS β/SE	IV β/SE
Announcements (mill. per cap)	**4.304**	**8.075**	—	—	—	—
	(2.372)	**(3.540)**				
Obligations (mill. per cap)	—	—	**6.388**	**10.196**	—	—
			(2.984)	**(3.599)**		
Payments (mill. per cap)	—	—	—	—	**15.045**	**22.163**
					(6.213)	**(8.622)**
Change in PI moving average	**−4.819**	**−4.519**	**−4.786**	**−4.562**	**−4.358**	**−3.978**
	(1.350)	**(1.306)**	**(1.330)**	**(1.263)**	**(1.342)**	**(1.308)**
Tax benefits (mill. per cap)	**13.215**	**18.270**	**8.829**	**9.656**	5.337	4.341
	(5.902)	**(6.711)**	**(4.946)**	**(4.698)**	(4.916)	(4.712)
Dec07–Feb09 trend	**0.258**	**0.194**	**0.264**	**0.224**	**0.297**	**0.280**
	(0.070)	**(0.082)**	**(0.065)**	**(0.066)**	**(0.058)**	**(0.057)**
Feb09 level	**−0.035**	**−0.040**	**−0.032**	**−0.033**	**−0.030**	**−0.030**
	(0.014)	**(0.014)**	**(0.014)**	**(0.013)**	**(0.014)**	**(0.013)**
2003–2007 house price growth	0.001	−0.002	0.003	0.001	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Constant	−0.001	−0.005	−0.003	−0.007	−0.002	−0.005
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
Observations	50	50	50	50	50	50
R^2	0.580	0.555	0.591	0.575	0.602	0.590
Robust first-stage F		9.448		21.708		11.651
Overidentifying restrictions test (p-value)		0.154		0.511		0.214

pre-ARRA employment trend in the state), the initial level of the employment per capita in February 2009, and the percentage change in the state's House Price Index from 2003 through 2007 (a measure of the pre-recession run-up in house prices). The stimulus variables are measured in millions of dollars per capita.

The first two columns of Table 5 show the results with stimulus measured by cumulative announcements through February 2010. The OLS estimate of the jobs multiplier, β, is 4.3, with a standard error of 2.4. As shown in equation (2), this number can be interpreted as saying that each $1 million of ARRA announced funds received by a state is associated with 4.3 jobs created or saved in that state (between February 2009 and February 2010). The IV estimate is 8.1 (s.e. = 3.5). This estimate implies that the ARRA spending's cost per job created or saved at its one-year mark was approximately $123,000. Using cumulative obligations as the stimulus measure, the OLS estimate is 6.4 (s.e. = 3.0) and the IV estimate is 10.2 (s.e. = 3.6). This IV estimate implies a cost per job of about $98,000. For payments, using either OLS or IV, the estimated multiplier is much less precisely estimated than for announcements or obligations, but it is still statistically significant. The OLS estimate is 15.0 and the IV estimate is 22.2. The latter implies a cost per job of about $45,000. The fact that in every case the OLS estimate is lower than the IV estimate suggests that stimulus spending may have been endogenously directed disproportionately toward states experiencing worse economic outcomes. In other words, OLS estimates of ARRA spending's impact are likely to be biased downward. Yet, it is worth noting that even the OLS estimates are found to be positive and statistically significant.

The first-stage F-statistics are shown at the bottom of the table. Stock and Yogo (2005) provide critical values of these statistics below which indicate possible weak instrument bias. At conventional significance levels, they list a critical value of 12.83 for the case of one endogenous variable and three instruments. The first-stage statistic for the obligations regression is well above that critical value. The statistics for the announcements and payments regressions, though, are slightly below the critical value, suggesting the coefficient on announcements may be biased downward (toward OLS). Also shown are the p-values corresponding to the Hansen (1982) J-test of overidentifying restrictions. The p-values in all cases are above conventional significance levels.

It is also worth mentioning the estimated coefficients on the control variables. I find that the 2005–2006 change in the three-year average of personal income is negatively and significantly associated with employment change over February 2009 to February 2010. This may reflect the fact that states that grew faster during the mid-2000s boom tended to experience larger economic declines during the 2007–2009 recession and its aftermath. Estimated ARRA tax benefits are found to have had a positive effect on employment; the effect is statistically significant when stimulus spending is measured by announcements or obligations but not when it is measured by payments.[20] The pre-ARRA trend (from December 2007 to February 2009) in employment change (per capita) is positively associated with the post-ARRA employment change; the coefficient is statistically significant in all cases.

---

[20]The coefficient on estimated tax benefits should be interpreted with caution. This variable is an estimate of the expected tax benefits the state will receive over the entire 10-year horizon of the ARRA rather than actual tax benefits from February 2009 to February 2010, which is unobserved.

This result likely reflects positive momentum or inertia in employment growth during this period. I also find that the initial level of employment in February 2009 is negatively associated with post-February 2009 employment change, suggesting some conditional convergence across states in terms of employment-population ratios. Lastly, the pre-recession run-up in house prices in a state has no significant effect on post-ARRA employment growth, conditional on the other covariates. This lack of effect could be because any negative effects of the housing run-up on subsequent employment growth is already picked up by the pre-ARRA employment trend variable.

Returning to the coefficients on the stimulus measures, there are clear differences in the magnitudes of the estimated jobs multiplier across the three stimulus measures. In particular, the multiplier based on payments—which is the typical measure of government spending in prior empirical work—is much higher than that based on announcements or obligations. How should these differences be interpreted? One interpretation is that announcements, obligations, and payments are just three alternative measures of a latent variable representing the true expected amount of stimulus received by each state from the ARRA. Announcements and obligations have much larger mean values (as shown in Table 3) than payments but all three measures are very highly correlated. Thus, to a first approximation, one can think of payments as simply proportional to obligations or announcements: $\text{payments}(i,t) = \alpha \times \text{obligations}(i,t)$, where $\alpha < 1$. If this were literally true, then the estimated ARRA spending effect based on payments would simply be $\alpha \times \beta$, where β is the estimated multiplier based on obligations. This appears to be approximately true empirically: the ratio of the average level of payments per capita to the average level of obligations per capita, as of February 2010, is 0.48 (see Table 3). The ratio of their coefficients in Table 5 is 0.46.

As discussed in Section IIC, announcements and obligations are more likely to reflect the magnitude of the true latent stimulus variable because they should lead the economic activity generated by the stimulus, whereas payments are likely to be a lagging indicator of announcements and obligations. For this reason, the coefficient on payments in these type of regressions is likely biased upward because it is reflecting the true effect stemming from announcements or obligations which won't be recorded in payments until several months later.

C. Robustness

In Table 6, I evaluate the robustness of the baseline results. First, I consider whether the results are sensitive to the presence or absence of the control variables included in the baseline specification. The second row of the table shows the IV coefficients on the stimulus variables when controls are omitted. The estimated effect of ARRA spending on employment is positive and significant regardless of whether the controls are included or not. The coefficient on announcements is quite similar to the baseline case. For obligations and payments, the coefficient without controls is even larger.

A possible concern one might have regarding the baseline results is that the economies of less populous states have tended to perform better during the 2007–2009 recession and subsequent recovery than other states. Even though it is difficult to

TABLE 6—IV REGRESSION RESULTS, ALTERNATIVE SPECIFICATIONS, AND DATA CHOICES
COEFFICIENT ON ARRA SPENDING MEASURE

	Announcements β/SE	Obligations β/SE	Payments β/SE
Baseline	**8.075**	**10.196**	**22.163**
	(3.540)	**(3.599)**	**(8.622)**
No controls	**7.258**	**15.688**	**48.743**
	(2.990)	**(4.010)**	**(12.443)**
Control for population	**10.797**	**11.438**	**21.832**
	(4.548)	**(3.672)**	**(7.573)**
Control for industry composition	**8.134**	**10.287**	**21.935**
	(3.785)	**(3.714)**	**(8.688)**
BLS sampling weights	**2.288**	**3.158**	**7.745**
	(1.531)	**(1.685)**	**(4.029)**
QCEW	**8.053**	**9.725**	**21.287**
	(3.461)	**(3.624)**	**(8.709)**
Drop outliers (Montana, South Dakota,	**19.386**	**14.243**	**22.702**
Alaska, North Dakota, Wyoming)	**(8.218)**	**(4.257)**	**(8.305)**
Sensitivity to instruments			
Drop ED instrument	**9.080**	**10.788**	**26.829**
	(3.690)	**(3.657)**	**(9.312)**
Drop HHS instrument	5.918	**10.409**	29.032
	(3.833)	**(5.780)**	(22.607)
Drop DOT instrument	8.378	**9.448**	**18.518**
	(6.617)	**(4.506)**	**(9.067)**
Sensitivity to choice of pretreatment period			
Jan 09	**9.838**	**12.053**	**27.914**
	(3.850)	**(3.737)**	**(9.942)**
Dec 08	8.798	**13.901**	**31.849**
	(5.459)	**(5.223)**	**(12.525)**
Nov 08	**10.009**	**14.439**	**34.724**
	(5.153)	**(5.091)**	**(12.817)**

think of why population would directly cause better economic outcomes, there is nonetheless concern that population may be correlated with one or more of the instruments, invalidating the instruments. Thus, as a robustness check, I report results in third row for regressions in which state population is added as an additional control variable along with the five included in the baseline specification. The estimated multipliers are similar or a bit higher than those from the baseline model.

Another concern might be that the ARRA formula factors could be correlated in some cases with state industry composition, which in turn may predict state employment change during 2009 and beyond. For example, state economies that are more reliant on agriculture or natural resources have tended to fare relatively better during the economic recovery and also tend to be more sparsely populated and thus have more lane-miles per capita, one of the factors determining ARRA grants from the DOT. Though the pre-ARRA trend in employment per capita should capture much of any such effects, it nonetheless presents a concern for instrument validity. Therefore, as a robustness check, I add to the baseline regression an industry shift-share variable, similar to Bartik (1991), that predicts state employment change (per capita) over February 2009 to February 2010 based on national employment change

over that period for each industry (NAICS supersector) and each industry's share of total state employment as of February 2009. The results are shown in the fourth row of Table 6. The inclusion of the industry shift-share variable has very little effect on the coefficient on predicted stimulus for all three measures of spending.

I next perform two robustness checks related to potential measurement error in the CES employment data. The first one addresses the concern that some states, especially less populous states, may have more measurement error in employment than others and should be given less weight in the regressions. The fifth row of the table presents results where states are weighted by the inverse of their sampling error variance from the CES payroll survey, as reported by the BLS. This weighting will also mitigate any undue influence of outlier states in terms of ARRA spending (such as Alaska, North Dakota, and Wyoming) because these sampling error variances are highly negatively correlated with state population. One can see that the multipliers obtained in the weighted regressions are generally quite similar to those obtained without weighting.

In the sixth row, I consider an alternative measure of state employment produced by the Bureau of Labor Statistics (BLS). The Quarterly Census of Employment and Wages (QCEW), previously known as the ES-202 series, is based on a census of state administrative (UI) records and thus has minimal measurement error. The estimated jobs multipliers based on QCEW data are very similar to those based on the CES data, suggesting that the CES-based results are not biased due to measurement error.

I then assess whether the baseline results are unduly influenced by outliers. As is apparent in Table 2 and Figure 4, five sparsely populated Western states—Montana, South Dakota, North Dakota, Alaska, and Wyoming—have notably higher values of both actual and predicted stimulus spending than other states. The seventh row of Table 6 shows the estimated jobs multipliers if one excludes these states from the sample. The multipliers are actually higher, especially for announcements, when one excludes these outliers.

Next, I investigate whether the baseline results are driven by any particular instrument. Rows 8–10 of Table 6 show how the IV coefficients change if one drops any single instrument. No single one of the instruments appears to be critical to obtaining the baseline results. The only exceptions are that the coefficients on predicted announcements and predicted payments are insignificant (with large point estimates and standard errors) when the HHS instrument is excluded and the coefficient on predicted announcements is large but insignificant when the DOT instrument is dropped.

In rows 11–13, I consider alternative choices for the initial, or pre-treatment, period. If both the passage and the geographic distribution of the ARRA was substantially anticipated prior to February 2009, then the Act may have had an economic impact prior to passage. In particular, such anticipation would imply that a state's February 2009 employment level is an inaccurate measure of "pre-stimulus" employment, and that using February as the initial month may underestimate the impact of the stimulus. Media reports and surveys from the time suggest that, while many agents likely anticipated a large fiscal stimulus package at least one or two months prior to the ARRA's enactment in February 2009, how those funds would be allocated across states was not widely anticipated prior to February. For example, even in the first few weeks after the ARRA's enactment, state officials complained that they had little notion of the amount

FIGURE 4. CHANGE IN EMPLOYMENT VERSUS PREDICTED ARRA ANNOUNCEMENTS (*in per capita terms*)

and composition of ARRA funds they would be receiving.[21] Nonetheless, I report here how the baseline results differ if one uses November 2008, December 2008, or January 2009 as the pre-treatment month in defining the pre-treatment to post-treatment employment change (as well as in defining the pre-stimulus employment level and trend control variables). In all cases, February 2010 is still the end month. Note that while employment in months prior to February 2009 are more likely to be unaffected by ARRA anticipation, using an earlier month also introduces more noise into the measurement of post-ARRA employment change.

The results are relatively insensitive to the exact month used as the pre-treatment period, though the standard errors are higher when an earlier month is used, consistent with the notion that using earlier initial months adds measurement error to the dependent variable. Overall, these results suggest that any understatement of the true impact of ARRA spending introduced by treating February 2009 as the pre-treatment month is relatively small, especially for the preferred measure of spending, announcements.

III. Extensions and a Falsification Test

A. *Employment Effects by Sector*

I turn now to two interesting extensions of the baseline results. First, I estimate the employment effects of ARRA spending on selected subsectors. Given large portions of the stimulus package were targeted at aid for state and local governments,

---

[21] The following quote from the spokesperson for the Governor of North Carolina, reported in the March 4, 2009 edition of the *Wall Street Journal* is a case in point: "All of the states are in the same boat—none of us have a very clear picture yet about the details of the package."

TABLE 7—IV REGRESSION RESULTS, BY INDUSTRY. COEFFCIENT ON ARRA SPENDING MEASURE

	Total nonfarm β/SE/F	Private nonfarm β/SE/F	S&L govt. β/SE/F	Construction β/SE/F	Manufacturing β/SE/F	Educ. and health β/SE/F
Announcements (mill. per cap)	**8.075** (3.540) *9.448*	3.965 (3.097) *10.253*	**3.005** (0.846) *8.130*	**3.590** (0.950) *12.941*	**0.820** (0.467) *20.060*	0.567 (0.387) *22.450*
Obligations (mill. per cap)	**10.196** (3.599) *21.708*	**8.188** (3.288) *21.159*	**2.415** (0.784) *25.395*	**3.324** (1.207) *23.032*	**1.112** (0.572) *33.099*	**1.036** (0.540) *23.784*
Payments (mill. per cap)	**22.163** (8.622) *11.651*	**20.660** (7.762) *11.566*	**5.771** (2.211) *10.039*	**7.091** (3.340) *10.890*	**3.171** (1.680) *12.170*	**3.953** (1.966) *6.254*

infrastructure, high-tech and green manufacturing, healthcare, and education, I look at the sectors of state and local government, construction, manufacturing, and private-sector education and health services.[22] Table 7 displays the estimated jobs multiplier for each of these subsectors as well as for the total nonfarm sector and the private nonfarm sector. (Note that these categories are not mutually exclusive; construction, manufacturing, and education and health services are three selected subsectors within the private nonfarm sector.) The regressions use the same set of instruments and include the same control variables as the regressions in Table 5, except that the initial level and pre-ARRA trend in employment per capita are defined in terms of the same subsector as the dependent variable. Results using announcements as the stimulus spending measure are shown in the top row; the second and third rows show the results based on obligations and payments, respectively. Standard errors are shown in parentheses, followed by the first-stage F-statistic in italics.

The jobs multiplier for the private nonfarm sector is estimated to be 4.0 for announcements, 8.2 for obligations, and 20.7 for payments.[23] The estimates are statistically significant for obligations and payments, but not for announcements. For the state and local government sector, the IV estimated jobs multiplier is positive and significant for all three measures of spending. The multiplier is found to be 3.0, 2.4, and 5.8 for announcements, obligations, and payments, respectively. The estimated jobs multiplier also is positive and significant for all measures of spending for the Construction and Manufacturing sectors. For the Education and Health sector, the multiplier is estimated to be positive and significant based on either obligations or payments, but not based on announcements.

B. *Alternative End Dates*

The cross-sectional analysis described above smoothes over any variation among states in the intertemporal pattern of stimulus spending and outcomes between the

---

[22] Unfortunately, employment data is not available for public-sector education and health services.

[23] Note that these jobs multipliers by subsector are expected to be smaller than the total nonfarm sector as long as the excluded subsectors have non-negative multipliers. Hence, their magnitudes are not directly comparable. In Section VIA, the magnitudes of the sector-specific multipliers will be evaluated as percentages relative to each sector's pre-ARRA level of employment, making them comparable across sectors.

ARRA's enactment and the end of the sample period. For example, for a given level of cumulative spending to date, one state may have received most of the spending early in the sample period whereas another may have received most of the spending later in the period. This timing variation may contain useful information, but it is likely to be endogenous for two reasons. First, as mentioned earlier, states with well-run governments may fulfill the requirements necessary to receive certain ARRA funds sooner than other states and having a well-run government may itself lead to better economic outcomes. Second, some components of the ARRA will be doled out to any given state in response to negative economic shocks as they hit the state, so again the timing of the stimulus will be endogenous with respect to the timing of economic outcomes. Unfortunately, while I arguably have strong and valid instruments for cumulative stimulus spending up to any particular post-ARRA-enactment date, I have no additional instruments that predict the *flow* of spending (i.e., the first-difference of cumulative spending) by month. Absent some exogenous determinant of the *monthly flow* of ARRA spending, the exogenous component of this monthly flow is unidentified. This rules out using a dynamic panel model to estimate a distributed lag structure or impulse response function for stimulus spending.

However, one can still assess how the employment effects of cumulative past ARRA spending varied over time—that is, over alternative sample end dates. How the estimated employment effect varies with the sample end date potentially reflects both the timing of stimulus spending up to the end date (i.e., did the bulk of spending occur toward the beginning of the sample or closer to the end date?) and the impulse response function for the spending. Again, it is not possible to disentangle the two without very strong exogeneity assumptions.

The panels in Figure 5 plot the estimated employment effect of ARRA spending for alternative end months, from as early as May 2009 to as late as March 2011. That is, each data point in the plots is the IV coefficient on ARRA spending from a separate regression of the baseline model, where the only difference across regressions is the end-month used to define the dependent variable, which is the change in employment from February 2009 to the end-month. The earliest month of data for announcements is August 2009. Panels A, B, and C show the results for announcements, obligations, and payments, respectively. In each panel, the solid line shows the estimated IV coefficient on cumulative ARRA spending (as of each month). The dashed lines indicate the 90 percent confidence interval. The dotted lines show the path of observed ARRA spending (in billions of dollars, indicated on the right axis) according to announcements, obligations, or payments.

Based on announcements, the estimated employment effect was positive and significant throughout the sample period (August 2009 to March 2011) except for one month, May 2010, in which the effect was positive but not statistically significant. Aside from the drop in May 2010, the magnitude of the employment impact grew over the sample period, rising from about 5 to about 12.[24] The pattern of coefficients

---

[24] Earlier drafts of this paper, based on real-time preliminary CES employment data from September 2009 onward (unlike this draft which uses the revised final CES data through September 2010 and preliminary data thereafter), showed a steep and sustained drop-off in the employment effects of ARRA spending starting around April of 2010. This drop-off goes away when one uses the revised data, suggesting the earlier result was likely an artifact of measurement error in the real-time data.
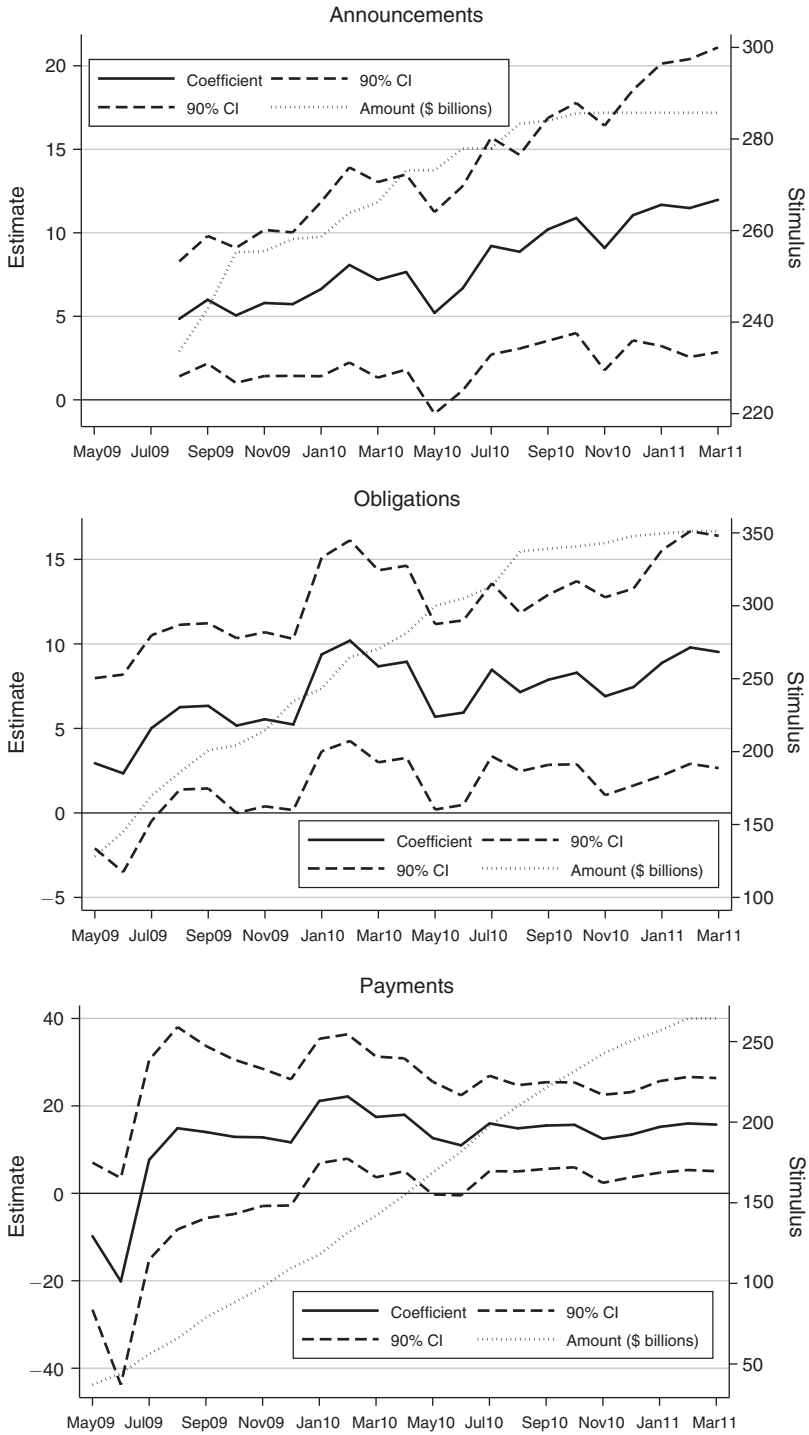
FIGURE 5. JOBS MULTIPLIER COEFFICIENT OVER TIME

using obligations is similar for the corresponding time period (August 2009 to March 2011). The coefficients based on payments are positive and significant from January 2010 through the end of the sample, except for in June 2010. However, unlike the pattern based on announcements or obligations, they do not exhibit an upward trend. Panels B and C also provide additional evidence for earlier months going back to May 2009. Based on obligations, ARRA spending did not begin to have a statistically significant impact on total nonfarm employment until August 2009. Based on payments, the confidence interval around the estimated employment effect is very large in the early months of the sample; the employment effect becomes statistically significant starting with the end month of January 2010.

C. A Falsification Test

To assess whether the ARRA employment effects estimated in this paper are unique to the post-ARRA-enactment period, and therefore less likely to be spurious, I perform a falsification test by running the following two-step estimation. In step 1, I estimate the baseline regression but without stimulus spending:

$$(4) \qquad (L_{i,T} - L_{i,0}) = \alpha_T + \mathbf{X}_{i,0}\Gamma_T + \varepsilon_{i,T},$$

where $\mathbf{X}_{i,0}$ is the same vector of pre-ARRA control variables as used in the baseline IV regressions. I estimate this model (by OLS) separately for each month from March 2008 through March 2011. One can think of this regression as providing a forecast (or backcast for months prior to February 2009) of employment in month T relative to February 2009.

In step 2, I regress the residuals from step 1 on a linear combination of the three instruments used in the baseline IV regressions. The coefficients on the instruments in this linear combination are simply their coefficients from the first-stage of the baseline IV regression as of February 2010 (i.e., those shown in Table 4).[25] In other words, this linear combination of instruments is simply predicted ARRA spending as of February 2010. If the positive relationship I find in the baseline IV regressions between predicted ARRA spending and post-February 2009 employment change is truly a causal effect, then there should be no correlation between these employment residuals and predicted ARRA spending prior to February 2009.

The estimated coefficient on predicted ARRA spending from this second step regression, and its 90 percent confidence interval, are shown in Figure 6. The estimated coefficient is near zero and statistically insignificant for all months up to January 2009. Aside from the correlation in this last pre-ARRA month, the lack of correlation for all earlier months indicates there is no general, spurious correlation between employment, conditional on the control variables, and the instruments. The negative and significant correlation in January 2009 could reflect some early anticipation effects prior to ARRA passage, as discussed in Section IVC above. Overall, the results of this falsification test bolster the case that the baseline IV coefficients,

---

[25] I use the first-stage coefficients from the obligations regression, but the results are similar if one uses the coefficients from the announcement or payments regressions.

Coefficient from Reduced Form Regression

Change in employment regressed on predicted ARRA spending and controls
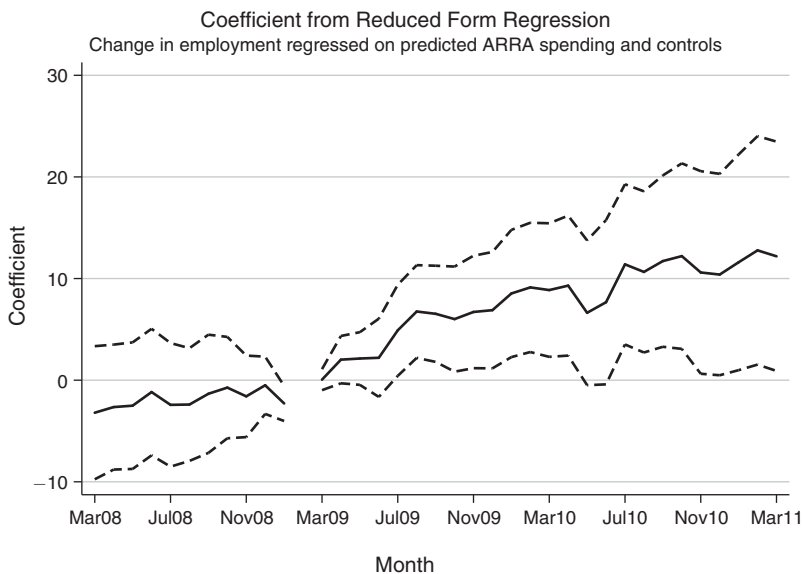

FIGURE 6. FALSIFICATION TEST

both at the one-year mark after the ARRA's enactment (Table 5) and over time (Figure 6), reflect the causal impact of ARRA spending on subsequent employment change and not a spurious correlation due to omitted factors. For omitted factors to explain the baseline results, the timing of their impact on employment would have to coincide very closely with the timing of ARRA spending.

IV. Implications and Comparisons with Other Studies

A. *Overall Impact of ARRA on National Employment*

The discussion of the results above focused on the sign and statistical significance of the estimated jobs multipliers. In this section, I turn to drawing out the economic implications of the results. As noted in the introduction, the jobs multipliers estimated in this paper are local multipliers. That is, they reflect the average employment effect within a state of ARRA spending received by that state. To extrapolate from these local effects to the national effect requires additional assumptions. For instance, if factor and goods mobility is greater among subnational regions than among countries, the local multiplier could be higher or lower than the national multiplier, depending on the sector (Moretti 2010). Additionally, if the government financing burden of stimulus spending falls on states independent of their ARRA receipts and agents are forward-looking, the local multiplier likely will be different than the national multiplier. That said, it is useful to consider the implications of the baseline multiplier estimates under the simple case of an economy with myopic agents and without trade.

In that case, one can calculate the nationwide number of jobs created or saved by ARRA spending implied by a given jobs multiplier estimate, by multiplying that

estimate by the amount of ARRA spending to date. Recall from Table 5, that the baseline IV regression for announcements indicated that each million dollars of ARRA announcements yielded 8.1 jobs in the total nonfarm sector by the end of the ARRA's first year. The analogous multipliers based on obligations and payments are 10.2, and 22.2, respectively. Announcements (excluding DOL) through February 2010 totaled about $260 billion. The jobs multiplier of 8.1 then implies 2.1 million more jobs in the economy in February 2010 than there would have been without the ARRA's spending. That number represents a 1.6 percent increase relative to the level of total nonfarm employment in February 2009. Analogous calculations using cumulative obligations or payments, and their respective jobs multiplier estimates, suggest nationwide effects of 2.6 or 2.8 million jobs, respectively. (Obligations and payments imply similar effects because the obligations' jobs multiplier is about half that of payments while cumulative obligations through February 2010 are about double cumulative payments).

Using the same ARRA spending total, one can calculate similar figures for the private nonfarm, state and local government, construction, manufacturing, and education and health services. The results based on announcements are shown in Table 8 below. The jobs multiplier estimate for private nonfarm implies 1.1 million jobs (1.0 percent relative to February 2009) created or saved as of February 2010. In the S&L government sector, ARRA spending is estimated to have created or saved 800 thousand jobs (5.1 percent) through February 2010. The construction sector is estimated to have created or saved 1.0 million jobs (16.4 percent) in the ARRA's first year, suggesting the sector was one of the key beneficiaries of the ARRA. Manufacturing is estimated to have increased net employment by 200 thousand jobs (1.9 percent) through February 2010. Lastly, I estimate that the education and health sector created or saved 200,000 jobs (0.8 percent) due to the ARRA through February 2010.

B. *Comparison with Other Studies*

How do these results compare to estimates from other studies of the number of jobs created or saved by the ARRA? I start with comparing it to the estimates from the most prominent and publicized governmental studies—the quarterly reports of the Council of Economic Advisers (CEA) and the Congressional Budget Office (CBO) (see CEA 2011 and CBO 2011). Both studies estimate the number of jobs created or saved due to total ARRA costs, including spending and tax cuts, by quarter. The reported ranges are shown in Table 9. In the top row, I provide the estimated number of jobs created or saved based on my estimated total nonfarm jobs multiplier multiplied by ARRA spending. The estimate here is based on announcements, my preferred measure of spending. (The estimates based on obligations or payments are slightly higher.) I show results for both the first quarter of 2010 and the first quarter of 2011, based on the jobs multiplier and spending totals as of March 2010 and March 2011, respectively. It should be noted, however, that the estimates for March 2011 are based on preliminary employment data that are subject to benchmark revisions. (The March 2010 estimates are based on revised, final data.) The ranges shown indicate the ranges over the three estimates yielded by the three alternative

TABLE 8—ESTIMATED NUMBER OF JOBS CREATED/SAVED BY ARRA SPENDING
(*in millions and percentages relative to February 2009*)

	February 2010	
Total nonfarm	2.1	(1.6%)
Private nonfarm	1.1	(1.0%)
S&L government	0.8	(5.1%)
Construction	1.0	(16.4%)
Manufacturing	0.2	(1.9%)
Education and health	0.2	(0.8%)

TABLE 9—ESTIMATED NUMBER OF JOBS CREATED/SAVED BY ARRA
(*total nonfarm sector*)

	First quarter 2010	First quarter 2011
This paper (spending only)	2.0 million	3.4 million
Congressional Budget Office	1.3–2.8 million	1.2–3.3 million
Council of Economic Advisers	2.2–2.6 million	2.4–3.6 million

spending measures (announcements, obligations, and payments). As of 2010:I, the CEA reports a range of 2.2 to 2.6 million jobs created or saved (see their table 8), whereas the CBO's range is 1.3 to 2.8 million jobs (see their table 1). As of 2011:I, the CEA's range is 2.4 to 3.6 million while the CBO's is 1.2 to 3.3 million. My estimate for 2010:I is roughly at the midpoint of the CBO range and is slightly below the CEA range. My estimate for 2011:I is slightly above the CEA range and toward the higher end of the CEA range.

Recall that the impact estimated in this paper relates only to ARRA spending, not ARRA tax reductions. If the ARRA tax reductions had a positive effect on employment (and the results shown in Table 5 for the coefficients on the 10-year estimated ARRA tax benefits suggest that they did), then the total ARRA (spending plus tax cuts) impact is even larger than indicated in the table. [26] (ARRA spending is about two-thirds of CBO-estimated ARRA costs through 2019.)

There are also a number of recent academic studies aimed at analyzing the overall effect of ARRA spending on employment. Similar to the approaches of the CEA and CBO, Blinder and Zandi (2010) use a large-scale macroeconometric model to simulate the economic effects of the ARRA and estimate that it increased employment (relative to the no-ARRA counterfactual) by the end of the first quarter of 2010 by 2.2 million jobs (see their table 7). By the end of 2010, they estimate the effect to be 2.9 million. (They don't provide quarterly estimates after 2010.) In sum, the number of jobs created or saved nationally based on the local jobs multipliers estimated in this paper, under the assumption of myopic agents and no cross-state spillovers, are generally similar to those generated by the macroeconometric modeling approach.

---

[26] Of course, there is much debate about whether tax cuts or spending have a larger fiscal multiplier. For studies addressing this issue, see Blanchard and Perotti (2002), Mountford and Uhlig (2008), Alesina and Ardagna (2010), and Barro and Redlick (2011).

Recent papers by Feyrer and Sacerdote (2011) and Conley and Dupor (2011) follow the cross-sectional IV methodology used in this paper. Feyrer and Sacerdote report results based both on county-level and state-level variation, focusing primarily on the latter. Using data on ARRA payments, they estimate a cost per job created or saved, as of October 2010, of $170,000 for the total nonfarm sector, which implies a jobs multiplier per million dollars of spending of about six. Conley and Dupor use a somewhat different measure of stimulus, dividing spending (measured by announcements or payments) by the pre-ARRA level of state tax revenues, based on the assumption that the effective amount of stimulus provided by $1 of ARRA spending is greater the lower the state government's pre-ARRA tax revenues (and hence the more fiscal strain the state is under). They find that, through September 2010, their measure of ARRA stimulus (based on obligations) increased S&L government employment but decreased employment in some private sectors, with essentially no net impact on total nonfarm employment.[27]

The estimated ARRA effects in both Feyrer and Sacerdote, and Conley and Dupor are much smaller than those found in this paper for the same time periods. As seen in Figure 5, I estimate a multiplier on outlays (payments) of about 16 as of September or October 2010. Even using announcements or obligations, I estimate a jobs multiplier for that time period that is near ten and statistically significant. There are various differences between the analysis done here and the analyses done in those papers, including differences in instruments, how spending is measured, and whether or not ARRA spending on extended unemployment insurance benefits are included. Another difference that could explain the divergent results is the vintage of the state employment data used. As of the time of this writing, the results in Feyrer and Sacerdote, and Conley and Dupor are based on preliminary, real-time payroll employment data (the initial BLS CES series) whereas the results presented here are based on the revised, final data. Earlier drafts of this paper, using the preliminary CES data, found a total nonfarm jobs multiplier as of October 2010 ranging from 0 using payments to three using announcements.

V. Conclusion

This paper analyzed the employment impacts of fiscal stimulus spending, using state-level data from the American Recovery and Reinvestment Act (ARRA) enacted in February 2009, and instrumenting for actual ARRA spending using pre-ARRA factors that went into the ARRA's allocation formulas. Cross-state IV results indicate that ARRA spending had a positive and statistically significant impact on total nonfarm employment at the one-year mark after the legislation was enacted. It also had a positive and significant impact on employment in the subsectors of state and local government, construction, manufacturing and, depending on which measure of stimulus spending one uses, the education and health sectors. Further analyses show that ARRA spending began having a statistically significant effect on total

---

[27] The decrease is statistically significant in their preferred specification for the combined sector of health, education, leisure and hospitality, and business and professional services, but is not significant for the goods-producing sector and is positive but insignificant for other service sectors (see their table 7).

employment around July or August of 2009, but not before. Moreover, there is no evidence of correlation between employment changes and predicted ARRA spending, conditional on controls, in months prior to the ARRA's enactment.

Based on my preferred measure of spending, announced funds, the results imply that its first year ARRA spending yielded about eight jobs per million dollars spent, or about $125,000 per job. Extrapolating from that marginal local effect to the national level, the estimates imply ARRA spending created or saved about 2.1 million jobs, or 1.6 percent of pre-ARRA total nonfarm employment, in that first year. The estimated employment effect is estimated to have grown further over time, reaching 3.4 million (based on announced funds) by March 2011. The estimates are moderately larger if one measures ARRA spending by obligated funds or actual outlays. Despite the use of a very different methodology, these estimates are in line with the range of estimates of the ARRA's impact generated by studies using the macroeconometric modeling approach.

It should be emphasized that the stimulus effects estimated in this paper correspond to the effects of one particular stimulus program enacted in a unique economic environment. There are at least two reasons why the ARRA's spending effects are not likely to be generally applicable to the issue of fiscal stimulus's effects in other contexts. First, the ARRA was unusual relative to previous stimulus programs in the United States and elsewhere in that it focused heavily on infrastructure spending and fiscal aid to state governments. These types of spending may well have different multipliers than federal government consumption expenditures. Second, the ARRA was enacted in a unique, and in many ways unprecedented, economic environment. The US economy was in the midst of its most severe economic downturns since the Great Depression. The resulting underutilization of resources could have made fiscal stimulus more effective than it would be in a more normal environment (see, for example, the results of Auerbach and Gorodnichenko (2012), finding that the fiscal multiplier is higher in downturns). Furthermore, monetary policy during the 2007–2009 recession and subsequent recovery was arguably stuck at the zero lower bound, or at least heavily constrained in its accommodative abilities. A number of theoretical studies have found that fiscal multipliers should be larger when monetary policy is less accommodative (e.g., Christiano, Eichenbaum, and Rebelo 2010).

Further empirical research is clearly warranted to study the extent to which the effects of fiscal spending along these dimensions. It is hoped that this paper serves as an illustration of how cross-regional variation in fiscal spending, combined with exogenous determinants of this variation, can provide useful reduced-form evidence on this question, and can serve as a complement to vector auto-regression and model-contingent approaches.

REFERENCES

Alesina, Alberto F., and Silvia Ardagna. 2010. "Large Changes in Fiscal Policy: Taxes Versus Spending." In *Tax Policy and the Economy*, Vol. 24, edited by Jeffrey R. Brown, 35–68. Chicago: University of Chicago Press.

Auerbach, Alan J., and Yuriy Gorodnichenko. 2012. "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4 (3): 1–27.

Barro, Robert J., and Charles J. Redlick. 2011. "Macroeconomic Effects from Government Purchases and Taxes." *Quarterly Journal of Economics* 126 (1): 51–102.

Bartik, Timothy J. 1991. *Who Benefits From State and Local Economic Development Policies?* Kalamazoo: William E. Upjohn Institute for Employment Research.

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 22 (1): 1–61.

Blanchard, Olivier Jean, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117 (4): 1329–68.

Blinder, Alan S., and Mark Zandi. 2010. "How the Great Recession Was Brought to an End." https://www.economy.com/mark-zandi/default.asp?src=economy_homepage.

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston. Forthcoming. "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy*.

Christiano, Lawrence J., Martin Eichenbaum, and Sergio Rebelo. 2010. "When Is the Government Spending Multiplier Large?" Federal Reserve Bank of Atlanta Center for Quantitative Economic Research (CQER) Working Paper 10-01.

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

Cogan, John F., Tobias Cwik, John B. Taylor, and Volker Wieland. 2009. "New Keynesian Versus Old Keynesian Government Spending Multipliers." Center for Economic Policy Research (CEPR) Discussion Paper 7236.

Congressional Budget Office. 2010. *The Budget and Economic Outlook: Fiscal Years 2010 to 2020.* US Congress. Washington, DC, January.

Congressional Budget Office. 2011. *Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output from October 2010 Through December 2010.* US Congress. Washington, DC, February.

Conley, Timothy G., and William Dupor. 2011. "The American Recovery and Reinvestment Act: Public Sector Jobs Saved, Private Sector Jobs Forestalled." Unpublished.

Council of Economic Advisers. 2010. *The Economic Impact of the American Recovery and Reinvestment Act of 2010, Second Quarterly Report.* Executive Office of the President of the United States. Washington, DC, January.

Council of Economic Advisers. 2011. *The Economic Impact of the American Recovery and Reinvestment Act of 2010, Fifth Quarterly Report.* Executive Office of the President of the United States. Washington, DC, November.

Drautzburg, Thorsten, and Harald Uhlig. 2010. "Fiscal Stimulus and Distortionary Taxation." National Bureau of Economic Research (NBER) Working Paper 17111.

Feyrer, James, and Bruce Sacerdote. 2011. "Did the Stimulus Stimulate? Real Time Estimates of the Effects of the American Recovery and Reinvestment Act." National Bureau of Economic Research (NBER) Working Paper 16759.

Fishback, Price V., and Valentina Kachanovskaya. 2010. "In Search of the Multiplier for Federal Spending in the States During the New Deal." National Bureau of Economic Research (NBER) Working Paper 16561.

Government Accountability Office. 2009. "GAO-10-223 Recovery Act: Recipient Reported Jobs Data Provide Some Insights into Use of Recovery Act Funding, but Data Quality and Reporting Issues Need Attention." Government Accountability Office (GAO). Washington, DC, November.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50 (4): 1029–54.

Ilzetzki, Ethan, Enrique G. Mendoza, and Carlos Végh. 2010. "How Big (Small) Are Fiscal Multipliers?" National Bureau of Economic Research (NBER) Working Paper 16479.

Leduc, Sylvain, and Daniel J. Wilson. Forthcoming. "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment." *National Bureau of Economic Research* (*NBER*) *Macroeconomics Annual 2012*.

Leeper, Eric M., Todd B. Walker, and Shu-Chun S. Yang. 2010. "Government Investment and Fiscal Stimulus." *Journal of Monetary Economics* 57 (8): 1000–12.

Mertens, Karel, and Morten O. Ravn. 2010. "Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach." *Economic Journal* 120 (544): 393–413.

Mian, Atif, and Amir Sufi. 2010. "The Effects of Fiscal Stimulus: Evidence from the 2009 'Cash for Clunkers' Program." National Bureau of Economic Research (NBER) Working Paper 16351.

Moretti, Enrico. 2010. "Local Multipliers." *American Economic Review* 100 (2): 373–77.

Mountford, Andrew, and Harald Uhlig. 2009. "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 24 (6): 960–92.

Nakamura, Emi, and Jon Steinsson. 2010. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." Unpublished.

Ramey, Valerie A. 2011a. "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 49 (3): 673–85.

Ramey, Valerie A. 2011b. "Identifying Government Spending Shocks: It's All in the Timing." *Quarterly Journal of Economics* 126 (1): 1–50.

Serrato, Juan C. S., and Phillippe Wingender. 2010. "Estimating Local Fiscal Multipliers." http://ceg. berkeley.edu/students_9_2583495006.pdf.

Shoag, Daniel. 2010. "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns." http://www.people.fas.harvard.edu/~shoag/papers.hmtl.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, edited by Donald W. K. Andrews, 80–108. New York: Cambridge University Press.

Wilson, Daniel J. 2012. "Fiscal Spending Jobs Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act: Dataset." *American Economic Journal: Economic Policy* http://dx.doi. org/10.1257/pol.4.3.251.