WHERE DID ALL THE MONEY GO? STIMULUS IN FACT AND FANTASY

Stephen A Marglin[1] and Peter M Spiegler[2]

Abstract

The Obama stimulus remains controversial even as we approach the fourth anniversary of its launch. The most thorough assessment of its impact, John Cogan and John Taylor's "What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package" (see also Taylor, "An Empirical Analysis of the Revival of Fiscal Activism in the 2000s") concludes that the stimulus had no impact on economic activity. Focusing on its impact on state governments, Cogan and Taylor contend that stimulus money simply allowed the states to build up their financial assets or reduce borrowing. We reassess the impact of the stimulus, focusing, like Cogan and Taylor, on the states. We find that the states spent about 2/3 of the stimulus money. Overall, we conclude that, over the period from mid-2009 to mid-2011 the stimulus added some 2 percent to GDP, in line with Congressional Budget Office estimates.

Our analysis has three parts. First, we analyze the regressions Cogan and Taylor interpret as supporting their contentions with regard to the impact of the stimulus on spending by the states. We find that these regressions do not support their conclusions. Based on aggregate time series of revenues and expenditures and including lagged dependent variables, the regression coefficients are misleading: because of serial correlation in the data, the regressions produce high coefficients on the lagged dependent variables and correspondingly low coefficients on the other variables regardless of whether the structure specified by Cogan and Taylor has any validity.

Second, we analyze the cross-sectional relationship between spending by state governments and injection of stimulus money. The data for Fiscal Year 2010 (July 2009 to June 2010) suggest that a dollar of stimulus money was divided between spending (2/3) and shoring up the state's balance sheet (1/3).

Third, we report the results of a survey of state budget officers. The results are remarkably uniform: despite differences in political orientation of their governments, and consequent differences in their evaluations of the wisdom of the stimulus, the general view is that the stimulus allowed the states to maintain expenditures that would have necessarily been cut in its absence.

July, 2013

[1] Professor of Economics, Harvard University, smarglin@harvard.edu

[2] Assistant Professor of Economics, University of Massachusetts-Boston, peter.spiegler@gmail.com

WHERE DID ALL THE MONEY GO? STIMULUS IN FACT AND FANTASY[3]

Stephen A Marglin and Peter M Spiegler

> It is of course possible that the planned surge in government spending will fail. Two or three years from now we could be facing a level of unemployment that is higher than today and that shows no sign of coming down. While it is too soon to examine in detail what might then be done, it is useful to consider the three possibilities. First, the level of government spending could be increased even more. To know whether this would help, it is important to study in detail the effectiveness of each of the different components of the spending surge. Martin S Feldstein, "Rethinking the Role of Fiscal Policy," January, 2009.

The 2009 American Recovery and Reinvestment Act was the most far-reaching experiment in fiscal stimulus in the history of the American economy. Like all fiscal stimulus programs, it was designed to raise GDP and employment above what they would have been in the absence of stimulus.

Did it work?

In fact the ARRA has been a Rorschach test of sorts, with assessments falling neatly along the left-right political divide: the left trumpets its success—or did until stimulus became a four-letter word spelled with eight letters—while the right regards it as a failed adventure in big government.[4] Even economists, who pride themselves on being responsive to, indeed, driven by, "the facts," have not been immune. Many have found in the ARRA confirmation for deep prejudices about the role of government, about conceptions of how the economy works, even

[3] Sam Harland provided excellent research assistance in transforming the raw data on stimulus grants into a form that could be analyzed by cross-section regressions. He contributed as well to formulating the regression equations reported in Section 3. We are not sure he knew what he was in for when he volunteered for this project, but he put in long hours, responding to many challenges that would have foiled a less resourceful and diligent analyst. Without Sam's help, we would still be sorting the data. Michael Ash provided important econometric advice. Noah Berger helped us to learn the ropes of state budgeting, as did Leslie Kirwan. Many government officials responded to our requests for information and clarification of data. Among these officials our chief debt is to the state budget officers and their staffs, but we are indebted as well to officials of the Bureau of Economic Analysis and the Census Bureau of the Department of Commerce, officials of the Department of Health and Human Services, and officials of the Board of Governors of the Federal Reserve System.

[4] The portion of the October 11, 2012 Vice-Presidential debate between Joe Biden and Paul Ryan devoted to the stimulus neatly summarizes the positions of the parties. The transcript is accessible at: http://www.debates.org/index.php?page=october-11-2012-the-biden-romney-vice-presidential-debate

about the methodology of economics. As we discuss below, there is not much more consensus among economists on this issue than there is among politicians, and the pros and cons tend to break down along party lines.[5]

Why has consensus been so hard to come by? Certainly, one reason is that the macroeconomy is complicated—there will almost always be *some* measure of controversy over the proper answer to a macroeconomic puzzle. But that has not been the central problem in this case. As we will argue below, the more important impediment has been a lack of clarity about the terms of the debate—the framework of assumptions underlying assessments of ARRA's effectiveness.

We can usefully distinguish three types of assumptions. *Identifying Assumptions* are those that allow us to interpret regression coefficients in the standard way—for example, in OLS, the assumption that the residuals are independent and identically distributed with mean zero and constant variance. *Behavioral Assumptions* are assumptions about the underlying behavior that generated the data—for example the assumption that agents engage in consumption smoothing. And *Counterfactual Assumptions* are assumptions about the values of the variables of interest in the state of the world in which the intervention had not occurred. Any valid assessment of ARRA (or any policy intervention, for that matter) must be built on a foundation of proper assumptions across these three areas: a properly identified empirical model, and a set of plausible counterfactual and behavioral assumptions used in interpreting the results of the analysis. For example, a regression that finds a positive effect of ARRA on GDP is evidence of ARRA's success only if the model used to generate the results is properly identified, one's counterfactual assumption is that GDP would have increased less than was observed, *and* one can provide sufficient evidence that the behavioral assumptions underlying the counterfactual assumption are plausible.[6] Although these standards may seem obvious, we argue that lack of clarity about them has muddled the debate over ARRA.

---

[5] See, for example, the exchange between John Cochrane (2010) and Joseph Stiglitz (2010). We do not mean to suggest that left and right-leaning economists who have cleaved to their party's line have done so solely for ideological reasons. We merely point out the tendency for the economic positions on ARRA to correlate with the political positions.

[6] These are only necessary conditions for the validity of the assessment, not sufficient conditions.

Our contention is that under empirically valid counterfactual and behavioral assumptions and a properly identified empirical model, the ARRA gets high marks.[7] We demonstrate this by reviewing the leading argument against ARRA's effectiveness—that rational agents smooth their consumption and, therefore, that temporary fiscal stimulus will be largely, if not completely, saved rather than spent—and showing that the premises and conclusions of the consumption-smoothing argument are highly sensitive to behavioral and counterfactual assumptions that are implausible in the case of the ARRA stimulus. By paying greater attention to the empirical realities of the execution of ARRA, we formulate more plausible behavioral and counterfactual assumptions and demonstrate that under these assumptions the evidence suggests that the ARRA stimulus achieved its stated goals.

We present our argument in four sections. The first section provides a brief overview of the debate over ARRA's performance, focusing specifically on the controversy over the value of the multiplier—the ratio of the total impact to the initial stimulus. In sections two through four, we expand on this thesis by undertaking a detailed critique of one particular application of the expenditure-smoothing argument: the striking claim that the stimulus failed to move the economy because not only individuals but also *institutional* recipients of Federal money and tax breaks—most notably, the state governments, who received roughly one third of the total ARRA outlays—did not spend any more than they would have in the absence of stimulus; rather they used the money to shore up their balance sheets. In this view the stimulus did nothing more than to provide debt relief—by substituting the debt of the US for the debt of the several states. These conclusions are founded on the position that the proper counterfactual is that states would have maintained spending at pre-recession levels in the absence of ARRA, and, therefore, that the ARRA can be judged a success only to the extent that the observed spending levels exceeded this counterfactual steady path.

In section 2, we review an analysis of the aggregate (national) time series data which appears to support the hypothesis that states used ARRA to shore up their balance sheets. We argue that the data do not support this conclusion because the analysis is plagued by invalid identifying assumptions—specifically, the assumption of no serial correlation. In section 3, we

[7] The question of whether or not ARRA was advisable from a policy perspective is a separate (and contentious) question, but, importantly, one that does not bear on the prior question of whether or not ARRA achieved its immediate aims relative to the proper counterfactual.

go behind the aggregate data to examine the differential impact of the ARRA money on spending across the 50 states. For this analysis, we operate under the working hypothesis that states *could have* borrowed in order to conduct business as usual with respect to spending and so might simply have substituted ARRA funds for withdrawals from their bank accounts. Our analysis leads to the conclusion that two-thirds of the money that went to the states was actually spent while only one-third went to shore up balance sheets. Finally, in section 4, we test the working hypothesis of Section 3, that, absent the ARRA, states could have borrowed enough to sustain pre-recession levels of spending. In this section we report the results of interviews with state budget officers about the impact of the ARRA on state government finances. The interviews were designed to elicit answers to the question of whether or not states could have borrowed as Section 3 (and Section 2) assume. These interviews provide a remarkably uniform set of responses that support a clear conclusion: with very few exceptions, the counterfactual claim that the states could have avoided spending cuts in the absence of the ARRA is implausible.

We conclude, in section 5, by discussing the general implications of our findings both for the ARRA and for the evaluation of fiscal stimulus programs in general. We argue that because of the politically charged nature of these programs and the attendant danger of ideological bias in their evaluation, it is especially important to hold the framing assumptions of any economic evaluation to a high standard of empirical fidelity. As we demonstrate here, direct appeal to the experience of the practitioners on the front lines of such programs can be an invaluable reality check on these assumptions, and a means of adjudicating disagreements among academic economists on this important issue.

1. General arguments for and against fiscal stimulus and their application to ARRA

The official website of the stimulus program, www.recovery.gov , divides the stimulus into three roughly equal parts, as in Table 1.

Table 1. The American Recovery and Reinvestment Act Writ Large	
$ Billions	
Tax Benefits	$ 297.8
Contracts, Grants, and Loans	$ 237.5
Entitlements	$ 229.2
Total	$ 764.5
Data from www.recovery.gov through July 13.2012	

Each of these categories includes a multitude of programs. Take "Tax Benefits." The largest single Tax Benefit program was an across-the-board benefit enjoyed by over 116 million taxpayers, Making Work Pay, which provided a $400 credit for an individual and a $800 credit for a couple with two working spouses for the 2009 and 2010 tax years, phasing out only at relatively high levels of $75,000 for an individual taxpayer and $150,000 for a couple filing joint returns. The total benefit amounted to $104.1 billion. At the other end of the spectrum were adjustments to the Alternative Minimum Income Tax totaling over $69 billion, which accrued to 13 million taxpayers. Tax Benefits also included some $11 billion of credits for improving residential energy efficiency, enjoyed by 6 million taxpayers and $33 billion of tax breaks for businesses.

"Grants, Contracts, and Loans" consisted chiefly of grants to states and local governments (the vast majority to states) for purposes ranging from education to highway construction and repair—$180 billion by the count of the Bureau of Economic Analysis, excluding Medicaid funding. But this category also included contracts with private entities totaling almost $30 billion, contracts which ranged from thousands of dollars to $1.5 billion awarded to Savannah River Nuclear Solutions to clean up the Savannah River Site, at which production of nuclear materials for the military's nuclear arsenal took place during the Cold War.

"Entitlements," the smallest of the three parts, consisted primarily of three programs: Medicaid grants to the states ($92 billion), extension of unemployment insurance ($61.8 billion), and family services ($41 billion, of which the lion's share was food stamps).

Controversy over the impact of the stimulus

The claim that recipients saved rather than spent stimulus money casts doubt on the whole idea of countercyclical fiscal policy. In the present political climate it is perhaps not surprising that politicians respond to countercyclical fiscal policy along party lines. Democrats credit the stimulus for keeping the economy going after the financial upheaval that culminated in the fall of Lehman Brothers in September of 2008. Many feared that the recession already under way for the better part of a year would turn into a depression of 1930s proportions, which hit bottom with one-third of the non-agricultural labor force without jobs. In contrast, US unemployment never climbed above 10 percent after Lehman went under, and the economy has been slowly improving since the middle of 2009, when the stimulus kicked in. Republicans argue that things would have gotten better much faster were it not for the Obama Administration's policies: the stimulus was a colossal waste which added neither jobs nor income, but only increased the national debt. Case in point: Solyndra, the solar panel manufacturer that received $527 million in ARRA guaranteed loans in 2009, only to go bust in 2011.

It is perhaps more surprising that economists are also divided. The Congressional Budget Office—whose estimates reflect (and perhaps to some extent mold) a consensus view—estimates that, at its peak in the spring and summer of 2010, the stimulus added between 700,000 and 3.5 million jobs, and between 0.8 and 4.6 percent to Gross Domestic Product (Congressional Budget Office 2012, "Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output from October 2011 Through December 2011," Table 1). These estimates, at least if we confine ourselves to the midpoints reported in column 7, are very close to what President Obama's advisers estimated the results of the stimulus would be, at least with regard to its incremental impact.[8] Table 2 summarizes the CBO's after-the-fact analysis and the estimates of the Obama economic team in advance of the stimulus.

Table 2. Impact of ARRA on GDP, As Estimated Before the Fact by Obama's Economic Team and Evaluated After the Fact by the CBO

		Before the Fact		After the Fact (All Estimates are as Percentages of GDP)		
(1)	(2)	(3)	**(4)**	(5)	(6)	**(7)**
Year and Quarter	GDP ($ billions SAAR)	Contribution of Stimulus to GDP ($ billions SAAR)	**As Percentage of GDP**	Low	High	**Midpoint**
2009 Q1	13893.7	31.2	**0.2**	0	0.1	**0.05**
2009 Q2	13854.1	75.8	**0.5**	0.4	1.4	**0.9**
2009 Q3	13920.5	194.8	**1.4**	0.6	2.4	**1.5**
2009 Q4	14087.4	227.6	**1.6**	0.7	3.4	**2.05**
2010 Q1	14277.9	265.5	**1.9**	0.9	4.3	**2.6**
2010 Q2	14467.8	313.5	**2.2**	0.8	4.6	**2.7**
2010 Q3	14605.5	330.7	**2.3**	0.7	4.1	**2.4**
2010 Q4	14755.0	336.0	**2.3**	0.6	3.5	**2.05**
2011 Q1	14867.8	307.1	**2.1**	0.6	3.2	**1.9**
2011 Q2	15012.8	226.4	**1.5**	0.4	2.5	**1.45**
2011 Q3	15176.1	195.4	**1.3**	0.3	2	**1.15**
2011 Q4	15319.4	169.3	**1.1**	0.2	1.5	**0.85**

Sources: St Louis Federal Reserve website for GDP; Bureau of Economic Analysis, Department of Commerce for primary effect of ARRA on components of GDP; Romer and Bernstein (2009) and authors' calculations for estimates in columns 3 and 4; Congressional Budget Office (2012) for estimates in columns 5, 6, and 7.

[8] Almost everybody inside and outside the Administration underestimated the severity of the downturn. Obama's economic team thus missed the mark with regard to *levels* of GDP and employment that the stimulus would achieve, but their predictions for the incremental impact of the stimulus were nonetheless close to the after-the-fact measurements of the CBO.

The range of its estimates runs from a lackluster impact at the low end (column 5)—never more than a 1 percent boost in GDP—and stellar at the high end (column 6), accounting for all the growth of the economy, and then some, in 2010. In the end, something for everybody. Rorschach wins.

While even at the low end of the estimates there is some stimulus to employment and output, there is a surprising amount of disagreement within the economics profession regarding the ARRA's contribution. According to a survey of 41 leading economists conducted by the Initiative on Global Markets (a project of the University of Chicago's Booth School of Business), only 80 percent of the respondents—the sum of those economists who "agree" and those who "strongly agree"—concurred with the view that the stimulus added jobs.[9]

Not only does 80 percent fall far short of the near-unanimity one might expect if economics lived up to its claims for scientific status, but the dissenters include many distinguished economists. Even before the stimulus was enacted, Harvard's Robert Barro pronounced it dead on arrival (2009). In his view the stimulus would crowd out other economic activity rather than unleashing a virtuous circle of spending. Chicago's John Cochrane (2009) argued that the stimulus could work only if enough of the people could be fooled into thinking the new debt would never have to be repaid. Stanford's John Taylor is perhaps the most widely known and the most vocal naysayer with regard to the stimulus, and Taylor—unlike Barro and Cochrane, who were writing before the ink was dry on the stimulus legislation—offers empirical evidence on the effects of the ARRA to make his case. Along with his Hoover Institution colleague, John Cogan, Taylor has forcefully challenged the consensus on the stimulus. Additionally, Taylor has made his views about the stimulus known not only in academia, but

[9] Regarding the makeup of the IGM panel, the organization's website explains:

> Our panel was chosen to include distinguished experts with a keen interest in public policy from the major areas of economics, to be geographically diverse, and to include Democrats, Republicans and Independents as well as older and younger scholars. The panel members are all senior faculty at the most elite research universities in the United States. The panel includes Nobel Laureates, John Bates Clark Medalists, fellows of the Econometric society, past Presidents of both the American Economics Association and American Finance Association, past Democratic and Republican members of the President's Council of Economics, and past and current editors of the leading journals in the profession. This selection process has the advantage of not only providing a set of panelists whose names will be familiar to other economists and the media, but also delivers a group with impeccable qualifications to speak on public policy matters.

also in more accessible form in the *Wall Street Journal*, on National Public Radio, and in congressional testimony. As he put it on NPR (August 14, 2011),

> I have looked at [the stimulus] with the numbers, looked at what happened, traced the money, and I don't find an impact. The studies that show it had an impact, they just simulate models. When I look at the data, where it went, temporary tax reductions went into people's pockets, they didn't spend it. This money that [was] sent to the states, they didn't spend it. They actually put it in their coffers. You can't see any impact on the infrastructure or the things that were supposed to happen. And those are the facts.

Taylor is on solid ground in pointing out that the CBO's evaluation presupposes a particular structure, a particular model of the economy, indeed the same model that served as the basis for Obama's advocacy of the stimulus. In fact, in arriving at its *ex post* assessment of the effectiveness of ARRA, the CBO does little more than to substitute the actual timeline of disbursements for the (*ex ante*) conjectural timeline of Federal disbursements that was used to argue for the stimulus in the first place. And though the details were modified as the legislation moved through the various committees of the House and Senate, the overall size of the stimulus did not change very much from the projected figure used by Obama's advisers to the figure actually enacted.

But Taylor is on less solid ground in claiming to be different from the analysts he criticizes—in effect claiming that he offers objective analysis while they offer ideology. According to Taylor, he and Cogan alone impose no model on the data but instead let the facts speak for themselves. "And those are the facts" he concludes on NPR. In fact Taylor's facts are interpretation, in the end no different from Christina Romer's or Robert Barro's—or ours. He, like everybody else, is arguing against the background of counterfactual assumptions—about what the states (or individuals) would have done without the stimulus—which in turn rest on a particular model of the economy; to justify this model he needs a plausible argument about behavior. There are no (counter)facts without a structure of interpretation.

As we turn to the assessments of ARRA's success, it is important to keep the centrality of interpretive frameworks in mind, as ultimately it is these frameworks—rather than the data themselves—that are at the core of the disagreement between supporters and detractors. We begin here with an examination of one of the central points of disagreement over ARRA: the controversy over the proper value of the multiplier. As we argue below, the differing positions on this issue stem mostly from disagreement about the underlying behavioral model. To an unfortunate extent, behavioral assumptions are kept in the background, even presented as self-evident features of the world. As a result the debate often devolves into two sides talking past each other and the issue seems to remain implacably partisan in nature.

Behind the controversy: of multipliers and men

The multiplier is a crucial element in the framework used by most analysts of the stimulus and has been a central element in the analysis of fiscal policy since Keynes—even though opponents of the stimulus would hardly agree that Keynes provides a preferred framework of analysis. Much of the disagreement between those who argue that ARRA positively impacted GDP and employment and those who argue it did not can be traced to different assumptions about the value of the multiplier and the multiplicand—what the multiplier was multiplying. Crowding out reduces the multiplier impact of each dollar spent because only a fraction of new spending, m, leads to new economic activity, the rest being offset by reductions in output and employment elsewhere in the economy. To take account of crowding out we modify the standard multiplier formula for expenditures, $\frac{1}{1-\mathrm{MPC}}$, by introducing m into both the numerator and denominator. The multiplier becomes $\frac{m}{1-m\mathrm{MPC}}$.

The disagreement about the multiplicand starts from the fact that most of the ARRA stimulus, as Cogan and Taylor (2012, pp 89-91) remind us, did not take the form of direct purchases of goods and services by the Federal government. Instead, the stimulus was largely focused on transfers and tax breaks to individuals and businesses, as well as grants to states to supplement the already massive grants-in-aid that have been part of our fiscal system for the past generation. This would not matter to the calculation of the multiplier if the beneficiaries of Federal largesse were themselves to spend all the stimulus money they receive or, in the case of tax breaks, all the money they didn't have to pay to the IRS. In this extreme case, the numerator of the multiplier formula would continue to be determined by the degree to which expenditure simply crowds out other production. The standard tax multiplier takes the more modest view that recipients of grants and tax breaks spend in the same proportion as the multiplier assumes for the representative agent, which leads to replacing the expenditure multiplier by the so-called tax multiplier $\frac{\mathrm{MPC}}{1-\mathrm{MPC}}$. The numerator becomes the marginal propensity to consume rather than 1. But suppose the direct beneficiaries—individuals, businesses, state governments—take a more conservative approach to spending. At the conservative extreme, if none of the original stimulus is spent, the multiplier would be zero. More generally, the tax multiplier becomes $\frac{v}{1-\mathrm{MPC}}$, where the parameter *v* represents the portion of the tax benefit, transfer payment, or grant actually spent by the original stimulus beneficiary.

Critics who argue that the stimulus did not stimulate in effect take extreme positions about the value of *m* or *v*. Barro's position with regard to crowding out, for example, assumes $m = 0$. Cochrane (2009) agrees with Barro but also argues that $v = 0$ because any rational agent who receives a tax cut, transfer, or grant will take into account the debt the Federal Government incurs to finance the stimulus. If she does her arithmetic, she will, according to a line of argument developed by Barro in the 1970s and 1980s (Barro 1989), put the stimulus money into a bank account to repay her share of the new taxes that will be required to pay off the debt. And of course the rational agents who are shut out from the stimulus will still recalculate their spending to take account of their future tax obligations. The result is a tie: according to the theory of Ricardian equivalence (the term of art for Cochrane's logic), the new spending by stimulus recipients will be just cancelled out by spending reductions elsewhere in the economy. Cogan and Taylor also argue for $v = 0$, but on the basis of expenditure smoothing along the lines of the permanent income and life cycle hypotheses (Friedman (1957; Modigliani and Brumberg, 1954; Ando and Modigliani, 1963).

Both crowding out and Ricardian equivalence seem to us relatively easy to refute. Crowding out was almost certainly a reality in World War II, when military spending quickly absorbed the margin of unused and underutilized resources that were the legacy of the Great Depression. But Barro is wrong to believe that this episode in our history has much bearing on the Great Recession: by mid-2009, when the stimulus kicked in, the unemployment rate had climbed to almost 10 percent.

Cochrane's endorsement of Ricardian equivalence appears to us to be grasping at straws in the hyper-rationality it imputes to agents: how many of us could do, much less actually *do*, the calculations implied in Ricardian equivalence? And then there is the factual assumption of the theory, namely, that the Federal Government will indeed repay its debt: even the deficit hawks concern themselves with whether or not the debt can be kept to a manageable proportion of

GDP, not with whether the debt will ever be fully repaid. It seems to us telling that Barro himself, the architect of Ricardian equivalence, did not see fit to invoke this line of thought in attacking the stimulus.

Cogan and Taylor's assumption of expenditure smoothing—and the corresponding counterfactual assumption that in the absence of ARRA agents would have temporarily depleted their bank accounts in order to maintain customary levels of expenditure—is harder to dismiss. A notable feature of Cogan and Taylor's argument is the idea that the same logic that applies to households also applies to state and local governments. Expenditure smoothing is standard fare in economics when it comes to households, but it is relatively novel to apply it to state and local government.[10]

For the purposes of this paper, however, the question is not merely whether consumption smoothing is a plausible behavioral assumption. Cogan and Taylor are likely right that consumption smoothing takes place. The question is by whom, how much, and over what time period. What fraction of households follow the dictates of Friedman and Modigliani? How extensively? Do households smooth their consumption over weeks (almost certainly), months, or decades? Does business engage in its own variety of expenditure smoothing when it comes to capital budgeting? To what extent do the states build up their bank accounts in good times and draw down these accounts in lean times as a buffer to maintain a stable trajectory of expenditures?

---

[10] Relatively novel but not totally so. Cogan and Taylor note Edward Gramlich's pioneering work on the effects of Federal grants on state budgets. Gramlich ends up skeptical of the efficacy of trying to stimulate the economy through grants to states, arguing as does Taylor a generation later, that grants end up fortifying state balance sheets (Gramlich 1978, 1979).

Before Gramlich, the terrain of how government spending is determined had been pretty much left to students of politics. As early as the 1960s, Aaron Wildavsky argued the position that would later inform Gramlich's work: last year's expenditures are the primary determinant of this year's expenditures. An important difference between Wildavsky on the one hand and Gramlich, Taylor, and other economists who invoke consumption smoothing is that Wildavsky claimed no rational basis—on the contrary—for the workings of the budgetary process. (See Wildavsky (1964) and Davis, Dempster and Wildavsky (1966, pp. 529-547; 1974, pp. 419-452). Nor did he apply his arguments to the operation of state and local government. His focus was rather on the process which determined agency budgets within the federal government, an altogether different environment from the states and cities, in which, for starters, no balanced budget constraints operate.

A comprehensive exploration of these questions would be beyond the scope of this paper. Two observations will have to suffice. First, many agents are simply unable to engage in expenditure smoothing. They have little or no saving and equally little access to credit markets. And the economics profession is well aware of such circumstances. There is indeed a large literature on what are called in the jargon "liquidity constrained households." Second, the economist's notion of "rationality" in "rational consumption smoothing" makes untenable demands on decision makers with respect both to their intertemporal utility functions and their needs and wants, much less about their future incomes, for the framework of the standard theory of consumer choice to make sense. (Most people, not everybody: the late James Duesenberry once quipped that the life cycle hypothesis is exactly the theory one would expect from a middle-aged college professor, thus demonstrating that some people's quips are as profound as other people's theories.) Instead, people fall back on habit, rules of thumb and other perhaps less elegant but more realistic ways of coping than that dictated by the economist's ideas of optimal planning (Marglin 2008, pp. 119-122). Moreover, real-world rationality may require people to put a higher premium on solidarity and sharing than the economist's paradigm of individual choice allows. A poor person embedded in community may feel that sharing a tax rebate with her less fortunate neighbors, particularly the neighbor faced with eviction if the rent goes unpaid or a blackout if paying the electricity bill is put off, is a higher priority than replenishing her own bank account. She knows that someday it will be her turn to rely on the community (Stack 1975, quoted in Marglin 2008, p. 23)

Our conclusion is that the circumstances of agents and the particulars of the stimulus package will affect the magnitude of v. Our behavioral assumption is that tax rebates and transfers directed towards lower income households are more likely to be spent, both because of liquidity constraints and "non-rational" behavior. This leads us to hypothesize that v is a negative function of household income. As far as state and local governments are concerned, we take liquidity constraints (in the form of balanced budget requirements) much more seriously than do Cogan and Taylor as limitations on expenditure smoothing. Our hypothesis is that in the absence of the ARRA, states would have had to drastically curtail expenditures or raise taxes.

The remainder of this section deploys a decomposition of the multiplier that turns on the value of v in order to account for the results claimed for the ARRA by the CBO—without the need for sophisticated models of the economy. The parameter v becomes the critical variable because it is relatively easy to eliminate the possibility of crowding out in the circumstances of the Great Recession, and so to fix $m = 1$, and because there is a wide consensus that during the kind of severe recession the economy experienced in 2009 and 2010 (and, we would say, the economy is still experiencing), the spending multiplier $\frac{m}{1\text{-}m\text{MPC}}$ reduces to $\frac{1}{1\text{-MPC}}$ and is of the order of 1.5.[11] Thus the tax (and transfer and grant) multiplier, given by $\frac{vm}{1\text{-}m\text{MPC}}$, becomes $1.5v$.

The stimulus in the light of spending propensities of beneficiaries

Consider again the three-part division of the ARRA into Tax Benefits; Contracts, Grants, and Loans; and Entitlements. A rough-and-ready division of Tax Benefits according to the specifics of the various programs makes it possible to identify eight programs that appear to benefit better-off segments of the population, listed in Table 3 along with the number of beneficiaries and the aggregate benefits of each. Table 3 also includes the two programs that benefit business.

Table 3. Tax Benefit Programs Chiefly Benefiting Better Off Taxpayers

Program	Nmber of Taxpayers Benefiting (Millions)	Benefits ($ billions)
Child Tax Credit	15.6	18.4
American Opprtunity Tax Credit	10.2	25.5
First Time Homebuyer Tax Credit	1	10.4
Extension of Alternative Minimum Tax Relief for Nonrefundable Personal Credits	13	4.3
Increased Alternative Minimum Tax Exemption	13	64.7
Residential Energy Credit	4.3	11
Business Incentives		33.5
Manufacturing and Economic Recovery		9.2
TOTAL		177

Sources: For description of programs and estimates of numbers of beneficiaries, http://www.recovery.gov/Transparency/fundingoverview/Documents/TaxBenefitInformation.pdf. For estimates of stimulus money received, http://www.recovery.gov/Transparency/fundingoverview/Pages/taxbenefits-details.aspx#IndividualTaxCredits.

We suppose that the rest—mainly Making Work Pay, the $100+ billion program that was the centerpiece of the ARRA tax breaks for individuals; the Earned Income Credit ($5.1 billion);

[11] Valerie Ramey provides a recent survey of the literature and concludes that "in a severe recession, the estimates are likely to be at the upper bound of [the] range [0.8 to 1.5]" (2011, p 681).

and the exclusion of a portion of unemployment benefits from taxable income ($6.3 billion)—roughly $120 billion in all, went to taxpayers who would not have engaged in expenditure smoothing to the extent that those higher up the income distribution did. For the purpose of modeling the impact of ARRA, we suppose recipients spent *all* this money, which is to say $v = 1$ for these programs.

With regard to the second category in Table 1—Contracts, Grants, and Loans, money flowing from the Federal government mainly to the states—our estimate (see Section 3 below) is that $v = 2/3$, that is, 2/3 was spent and 1/3 went to shore up state government balance sheets. For reasons that will be elaborated, we include the roughly $90 billion of Medicaid grants with other money going to the states. Applying the 2/3-1/3 split to all the money in the Contracts, Grants, and Loans portion of the ARRA, we calculate that 2/3 x $(240 + 90) = $220 billion was actually spent.

This leaves the remaining Entitlement portion. We assume this was money was spent by the immediate beneficiaries, that is, $v = 1$. Going as it did largely to recipients of extremely limited means—means in the sense both of the buffer needed to engage in consumption smoothing and the means to plan spending in terms of the long horizons assumed by Friedman and Modigliani—it is hard to see how much of it would have gone into recipients' bank accounts. Leaving out the Medicaid portion (already factored into Contracts, Grants, and Loans), we calculate spending as $230 – $90 = $140 billion, or more precisely, $137.2 billion.

The resulting spending from the three components of the Recovery Act stimulus is reported in Column 3 of Table 4.

Table 4. Reported Spending, Direct Spending, and Total Impact of the ARRA

$ Billions

	1	2	3.00	4	5	6
Program	Recovery.gov	Reallocating Medicaid to Grants	v	Direct Spending of Recipients = v x col (2)	Multiplier	Total Impact = col (4) x col (5)
Tax Benefits to upper income groups and businesses	177.8	177.8	0.00	0.0	0.0	0.0
Tax Benefits to lower income groups	120.0	120.0	1.00	120.0	1.5	180.0
Contracts, Grants, and Loans	237.5	329.5	0.67	220.8	1.0	220.8
Entitlements	229.2	137.2	1.00	137.2	1.5	205.8
Total	764.5	764.5		478.0		606.6

Data for Column 1 from www.recovery.gov through July 13.2012. Remaining data, authors' calculations

If we take the stimulus as being spent over the two years beginning in mid-2009 and ending in mid-2011, we have a total impact just over $300 billion per year, or roughly 2% of GDP, more or less in line with the CBO, but on the basis of a simpler and more transparent model which we believe reveals more clearly the critical assumptions underlying the analysis.

This calculation highlights—as Cogan and Taylor emphasize—that a key issue is how much of the stimulus actually got spent: the numerator of the multiplier formula. Most of the variability of multiplier, as the CBO multiplier estimates themselves show (see Congressional Budget Office 2012, Table 2, pp 6-7), depends on the numerator, on how much of the initial injection of stimulus is actually spent. In fact, the MPC, as has been noted, seems relatively uncontroversial, with a variety of estimates clustered around 1/3, as indicated by multipliers clustered around 1.5 (Ramey, 2011, p 681).

Cogan and Taylor are right to refocus the discussion on what is getting multiplied—even if our analysis leads us to very different conclusions about the size of the critical parameter, v. There is an important corollary. If the point is to stimulate the economy, it's necessary to put money in the hands of people who will spend it, a consideration which speaks in favor of targeting tax breaks, rebates, etc. towards low income recipients. It is hard to imagine that tweaking the alternative minimum tax is going to lead to considerable spending, despite the fact that it reaches further down the income ladder every year. This and similar concessions to the well-off may be politically necessary in a system dominated by special interests, but these elements of the stimulus package should be considered as the political price to be paid for stimulus rather than as part of the stimulus itself. The concentration of tax breaks in the hands of the relatively well-off—60 percent by our reckoning—may also explain why Taylor could not find any statistical relationship between aggregate consumption and the ARRA tax benefits (Taylor 2011a, pp 688-692).

2. Pitfalls of Aggregate Time Series: Some Problems with What Cogan and Taylor think the States did with the ARRA Money

In contrast with our admittedly cursory examination of the data on tax credits and transfers to individuals and businesses, the heavy lifting of this paper is substantiating our assertion that the states likely spent 2/3 of the grants they received from the Federal government.

Cogan and Taylor rely on time series evidence to support the contrary view that v = 0, but this section demonstrates why that reliance is misplaced.

Cogan and Taylor point to budget-stabilization funds in support of their view that the states would have run up debts or run down bank accounts in the absence of the ARRA. Budget-stabilization accounts, or rainy-day funds as they are often called, were created over the last decades by all but 3 of the states (Arkansas only in 2010) precisely to insulate expenditures against fluctuations in revenues, against the vagaries of the private economy.[12] The ups and downs of the private sector have an immediate and strong impact on state income and sales tax collections, on which states collectively relied for almost 50 percent of revenues, not counting Federal grants, raised in fiscal year 2010, the last year for which detailed data exist.

States made extensive use of rainy-day funds in the current downturn, and at first glance it would appear from the aggregate amounts left in these funds that the states, as Cogan and Taylor contend, could have continued to draw down these funds even more if the ARRA had not come to the rescue. Between FY2008 and FY2009 the states reduced the size of the aggregate rainy-day fund from $33 billion to $29 billion, and in FY2010 to $21 billion. And in FY2011 the aggregate rainy-day fund actually increased. However the aggregate data hide the fact that two states, Alaska and Texas, which have access to oil revenues to stabilize their budgets, accounted for over $10 billion of the total in 2008 and actually increased their rainy-day funds to over $17 billion in FY2010, recession or no recession. Table 5 gives rainy-day fund figures with and without these two states.

Table 5. Rainy-Day Fund Balances by Fiscal Years, $ millions

	2008	2009	2010	2011
Rainy Day Fund Balances, All States	32,943	29,006	21,034	24,154
Alaska	5,601	8,898	10,364	11,065
Texas	4,355	6,276	7,693	5,041
Rainy Day Fund Balances, All Other States	22,987	13,832	2,977	8,048
Number of States With Zero Balance	5	10	15	12

Source: National Association of State Budget Officers, *The Fiscal Survey of States,* various years. Data for FY2011 are preliminary.

By the end of FY2009, almost 40 percent of the rainy-day funds of the other 48 states had been used up, and by 2010 almost 90 percent. The number of states with a zero balance in their

[12] Arkansas is still listed in the publications of the National Association of State Budget Officers as one of the exceptions, but according to the website http://law.justia.com/codes/arkansas/2010/title-19/chapter-6/subchapter-4/19-6-486/ it created a rainy day fund in 2010.

rainy-day funds went from 5 to 10 and then to 15 over the two year period. (The figures include the four states without rainy-day funds in 2008.) Only in FY2011 did rainy-day funds begin to recover. So while states may in principle be committed to expenditure smoothing, rainy-day funds were woefully inadequate when push came to shove. In any case rainy-day funds pale into insignificance in size compared with the $120 billion of ARRA grants the states received over FY2009 and FY 2010, not to mention the $125 billion received in FY2011: if spending could have been supported by drawing down assets, it was not assets in rainy-day funds!

Taylor (2011a) and Cogan and Taylor (2012) do not explicitly consider the paucity of rainy-day funds. Rather, they simply make the counterfactual assumption that there would have been no restrictions to states' capacity to smooth expenditure if they had faced the recession without ARRA funds; they do not specify where the states would have found the funds to permit expenditure smoothing.[13] As we will see below, this assumption plays a major role in their econometric analysis and is empirically unwarranted.

Their primary empirical strategy is based on time series regression of state and local government purchases of goods and services (G) and transfer payments (E) on non-ARRA revenues (R) and the ARRA stimulus (A), along with the lagged dependent variable (G_{-1} and E_{-1}). The results, the authors claim, suggest a positive effect of ARRA on transfers, but a negative effect on purchases of goods and services, and the negative effect offsets the positive one. The conclusion is that the ARRA ended up simply improving the balance sheets of state and local governments. The basic regression equations are

$$G = a_0 + a_1G_{-1} + a_2R + a_3A + \mu \qquad (1)$$

$$E = b_0 + b_1E_{-1} + b_2R + b_3A + \xi \qquad (2)$$

The results of these regressions are presented in the first three columns of Table 6, column 3 differing from column 2 only by the absence of the insignificant constant term.

[13] Taylor ignores the relatively inflexible barrier between current and capital expenditures (see below). In a debate on the impact of the stimulus (at Harvard University, February 28, 2012), Larry Summers observed that constitutional or statutory balanced-budget requirements prevented the states from borrowing and smoothing expenditure. Taylor, in response, suggested that states could borrow on capital account to adjust their overall spending. In response to an email request for clarification by one of the authors of this paper, Taylor repeated that "borrowing for infrastructure investment is one means of flexibility." (Taylor, personal communication, February 29, 2012.)

Limiting the analysis to current revenues and expenditures (i.e. omitting capital expenditures), we find that in the Cogan-Taylor model the negative effect of grants to state and local governments actually exceeds the positive effect on transfers.[14] The fourth column represents the calculated impact on the consolidated balance sheet of state and local governments. It is obtained from columns 1 and 3 (or columns 1 and 2) via the identity linking the variables, namely

$$L = R + A - G - E,$$

where

L = Net saving = Change in the consolidated balance sheet of state and local governments,

from which it follows that

$$L = -(a_0 + b_0) - a_1 G_{-1} - b_1 E_{-1} + (1 - a_2 - b_2)R + (1 - a_3 - b_3)A - \mu - \xi \tag{3}$$

Thus the coefficients reported in column 4 do not represent an independent regression, but rather fall out of the budget identities that link spending and revenue to net saving on a NIPA basis. There is no new information in column 4.

As noted, the coefficients in the first two columns of Table 6 differ from Taylor's own coefficients, but the results are qualitatively the same. If you subtract the negative effect of a \$1.00 grant on state government purchases of goods and services, namely, \$0.46, from the positive effect of \$1.00 of stimulus on transfer payments, \$0.25 (in column 3, without a constant

[14] Professor Taylor graciously made their regression data available to us, but for various reasons we felt it necessary to work with data directly derived from the NIPA and other sources. Apart from data additions and modifications subsequent to the time when Cogan and Taylor undertook their analysis (early summer, 2011), there are four reasons why our data differ and hence why our results differ from theirs.

First, we focus on current account whereas Cogan and Taylor lumps current and capital account figures together. Second, Cogan and Taylor include an imputation made by the Bureau of Economic Analysis that we exclude: the NIPA definition of the current component of their G includes capital consumption as an approximation to the value of services rendered by the stock of physical capital owned by state and local governments. Third, the Cogan-Taylor regressions are based on dollar amounts, which introduce a trend in the data. We partially eliminate this trend by expressing revenues and expenditures as fractions of potential GDP in Table 7. Finally, we remove what we regard as a largely spurious multiplication of observations by replacing their quarterly data with state fiscal year data—the fiscal year is the unit of action for state budgets and though there are occasional mid-course corrections, modifications of expenditures and taxes in the process of a single fiscal year, these are comparatively rare and the main effect of using quarterly data is to introduce a multiplicity of non-independent observations and to multiply problems of serial correlation.

term; \$0.26 in column 2, with a constant term) you are left with an overall impact of -\$0.21 on government spending as a whole. The coefficient has the wrong sign *and* differs significantly from zero.

Table 6. Regressions of Nominal Purchases and Transfers on Lagged Dependent Variables, Revenues, and ARRA

	1	2	3	4	5	6
	Eqn 1	Eqn 2	Eqn 2	Eqn 3	Eqn 4	Eqn 4
Dependent Variables	G	E	E	L	O	O
Constant	9.098	-1.212		-9.098	5.511	
	1.545	2.005			2.780	
G_{-1}	0.760			-0.760		
	0.071					
E_{-1}		0.859	0.886	-0.886		
		0.075	0.045			
O_{-1}					0.776	0.727
					0.082	0.067
R	0.169	0.054	0.046	0.785	0.237	0.284
	0.042	0.021	0.012		0.072	0.057
A	-0.462	0.264	0.245	1.217	-0.144	-0.092
	0.087	0.099	0.089		0.180	0.172
Adj R^2	1.000	0.999	1.000		1.000	1.000
Frequency/Time	FY/1969-2011	FY/1969-2011	FY/1969-2011		FY/1969-2011	FY/1969-2011
N	43	43	43		43	43

Newey-West standard errors appear below coefficients

The Cogan-Taylor interpretive structure, then, actually leads to a stronger conclusion than that of conservative conventional wisdom: insofar as the grants to state and local governments are concerned, the stimulus was actually *counter*productive rather than merely useless. Cogan and Taylor focus on the effect of the ARRA money on purchases of goods and services by state and local governments, so the favorable impact of the stimulus on transfers, even using their original numbers doesn't help to make up for the unfavorable impact on expenditures : "It is important," they say,

> to distinguish between two types of grants. First are those that state and local governments may directly use to finance purchases of goods and services. Grants for transportation projects and elementary and secondary schools are included in this category. The second type is transfers that supplement household resources. Federal Medicaid grants to states fall into this category. Under NIPA accounting conventions,

> state Medicaid expenditures are treated as transfer payments to households which raise their disposable personal income. Their impact on GDP depends on how much of the rise in income results in a rise in personal consumption expenditures. In addition, to the extent that higher federal Medicaid grants are fungible at the state level they may free up other state revenues, and their impact may also be reflected by higher state government purchases of goods and services (Cogan and Taylor, p 91)

The action for Taylor is in the first type of grant: "from a Keynesian stimulus perspective, the purpose of… sending grants to state and local governments is to get these governments to increase *purchases*" (Taylor 2011a, p 692, emphasis added). The justification for this focus on purchases is the behavioral assumption that agents smooth expenditure. Under this assumption, Federal transfers to households (tax breaks, one-time supplements to social security, and the like) had no impact on consumer spending, and there is no reason to treat transfers made by states and localities any differently.

But there is. Quite apart from whether or not all personal transfers can be lumped together—we think not—it makes no sense to consider transfers payments made by states in the same way that we look at direct transfers from the Federal government to individuals. Most direct Federal transfers to individuals come with few or no strings attached—think social security—and it is reasonable to consider such transfers simply as putting more money in the pockets of recipients. However, the bulk of transfer payments made by states and localities are not really payments to the nominal recipients except by NIPA convention. Medicaid, the largest single transfer program, appears in the national income accounts as a transfer payment to individuals, but the individual never sees any cash. The payments are actually made to vendors of medical goods and services—for visits to doctors, surgical procedures, prescription drugs—and are purchases of goods and services every bit as much as direct purchases by state governments. From what Taylor terms a Keynesian stimulus perspective, or from any other perspective, it makes little difference as to whether states purchase goods and services directly or purchase goods and services by making payments to vendors of medical services, pharmaceuticals, and medical devices.

But to make their point Cogan and Taylor don't need to distinguish transfers from purchases. If we run their basic regression with total current outlays (O) as the dependent variable, that is, without distinguishing between purchases and transfers, we obtain the results in columns 5 and 6 of Table 6. The coefficient on the ARRA variable is −0.14 in column 5 and −0.09 in column 6, and the t-statistic is barely 1.0 in column 5 and less than 1.0 in column 6. Once again, the ARRA shows itself to be ineffective in stimulating state spending, and it becomes moot what the multiplier would have been had the states spent their ARRA grants rather than pocketing the cash.

Normalizing the variables by dividing them each by potential (nominal) GDP removes some of the problems of using the trend-dominated data in its raw form. And the basic results do not change. The negative hit to purchases remains significant, both statistically and economically, and is not offset by the positive stimulus to transfers. The regressions on expenditure in columns 5 and 6 suggest a positive impact on spending overall—interestingly, the coefficient on A is indistinguishable from the coefficient on R—but the coefficient on A is statistically insignificant.

Table 7. Regressions of Normalized Purchases and Transfers on Lagged Dependent Variables, Revenues, and ARRA

	1.000	2.000	3	4	5	6
	Eqn 1	Eqn 2	Eqn 2	Eqn 3	Eqn 4	Eqn 4
Dependent Variables	G	E	E	L	O	O
Constant	0.016	0.003		-0.016	-0.001	
	0.004	0.003			0.005	
G_{-1}	0.594			-0.594		
	0.861					
E_{-1}		0.975	0.949	-0.949		
		0.041	0.031			
O_{-1}					0.736	0.739
					0.068	0.0664
R	0.134	-0.014	0.016	0.85	0.244	0.237
	0.046	0.028	0.008		0.074	0.0585
A	-0.357	0.239	0.286	1.071	0.276	0.269
	0.057	0.092	0.082		0.168	0.1616
Adj R^2	0.883	0.982	0.999		0.969	1.000
Frequency/Time	FY/1969-2011	FY/1969-2011	FY/1969-2011		FY/1969-2011	FY/1969-2011
N	43	43	43		43	43
Newey-West standard errors appear below coefficients						

How Statistics Can Lie Without Even Trying

Even without the (unwarranted, in our opinion) distinction between expenditures and transfers, then, the aggregate time series regressions seem to support the view that ARRA did not significantly increase state government spending. There is however a problem: the statistical evidence speaks with a forked tongue. The standard interpretation of the regression estimates reported in Table 7, would be that spending depends mostly on previous spending and relatively little on current revenues. But, as with any econometric result, the standard interpretation is valid only if the relevant identifying assumptions hold. For time-series regression, it is crucial that the variables not exhibit excessive serial correlation (or, that the serial correlation is appropriately corrected for). The presence of serial correlation can lead to spurious results if the regression coefficients are interpreted as though the equation had been properly specified.

In the case of the Cogan and Taylor regression, serial correlation in the data could generate the observed results—that lagged expenditure matters a lot and current revenue not much—even in the case where lagged expenditure is actually irrelevant to current expenditure.[15] Assume that current expenditure is actually orthogonal to lagged expenditure and depends only on current revenues. The correct specification of the process relating R and O would then be

$$O_t = a_0 + a_1 R_t + \mu_t \qquad (4)$$

where μ is the error term. And, also by assumption, the independent variable and the error term are serially correlated according to a first order process:

$$R_t = \rho_1 R_{t-1} + \xi_t \qquad (5)$$

$$\mu_t = \rho_2 \mu_{t-1} + \varepsilon_t \qquad (6)$$

It follows that

$$\mu_t = O_t - a_0 - a_1 R_t = \rho_2 \mu_{t-1} + \varepsilon_t = \rho_2 (O_{t-1} - a_0 - a_1 R_{t-1}) + \varepsilon_t$$

and, since

$$R_{t-1} = (\rho_1)^{-1} R_t - (\rho_1)^{-1} \xi_t$$

we have

---

[15] Here is the appropriate place to report that everything we know about the problem of spurious correlation in regressions with lagged dependent variables—apart from an initial suspicion that all is not well in the state of regression interpretation in such cases—we learned from an unpublished paper by Christopher Achen (2001), with a little bit of help from a paper by Luke Keele and Nathan Kelly (2006).

$$O_t = (1 - \rho_2)a_0 + (\rho_1 - \rho_2)(\rho_1)^{-1} a_1 R_t + \rho_2 O_{t-1} + \varepsilon_t + \rho_2(\rho_1)^{-1} a_1 \xi_t \qquad (7)$$

In short, the lagged dependent variable O_{t-1} sneaks in because of the serial correlation of the error term in the original equation and—when $\rho_1 > \rho_2$ and ρ_1 is close to unity—at the same time reduces the coefficient on the true explanatory variable R_t. In the case at hand, $\rho_1 = 1.006$ and $\rho_2 = 0.663$. The estimates of the constant term and the coefficient of revenues in Eqn 4 are $\hat{a}_0 = -0.024$ and $\hat{a}_1 = 1.079$, with the result that the implied coefficients in Eqn 7 are, as in column 5, $(1 - \rho_2)a_0 = -0.008$, $(\rho_1 - \rho_2)(\rho_1)^{-1} a_1 = 0.368$, $\rho_2 = 0.663$. Observe that these numbers are close to what is reported in column 1 as the results of estimating Eqn 7 by ordinary least squares,

$$O_t = b_0 + b_1 R_t + b_2 O_{t-1} + \nu_t \qquad (8)$$

for which, $\hat{b}_0 = -0.000$, $\hat{b}_1 = 0.238$, and $\hat{b}_2 = 0.742$. According to Achen (2001, pp 5-6), the formulas for the limiting values of the direct estimates of the coefficients in Eqn 7 are

$$\text{plim } \hat{b}_1 = \left(1 - \rho_1 \rho_2 \frac{1-R^2}{1-\rho_1^2 R^2}\right) a_1$$

$$\text{plim } \hat{b}_2 = \rho_2 \frac{1-R^2}{1-\rho_1^2 R^2}$$

where R^2 refers to Eqn 4, reported in column 2. Without knowing the true value of a_1 we cannot estimate the limiting value of b_1 from Eqn 4. But we can estimate this limiting value conditional on the estimate $\hat{a}_1$ in Eqn 4. On this basis we have plim $\hat{b}_1 = 0.304$. In addition we have plim $\hat{b}_2 = 0.714$.

Clearly, these results are consistent with the hypothesis that only revenues matter for expenditures and the regression coefficients that emerge from the specification in column 4 with lagged expenditures are spurious. This is not to say that the specification in column 1, in which the driving force is the lagged dependent variable, is without merit. It rather says that the impressive statistics that characterize this equation turn out to add nothing to the argument for this specification.

The path of wisdom would seem to be to go beyond aggregate time series in the search for data that might shed light on how state revenues and expenditures responded to the injection of ARRA money. The rest of this paper reports on two such investigations, one with cross-sectional data from the states for FY2010, the other a series of interviews with state budget officers. Both of these inquiries suggest that the states responded quite positively to the ARRA, even when they did so holding their noses.

3. Cross Sectional Analysis of the ARRA and State Spending

In this section we deploy cross-sectional evidence to test the hypothesis that states spent the bulk of the ARRA monies they received against the hypothesis that these monies had little or no effect on spending, instead going to shore up their balance sheets. This exercise provisionally commits us to the stipulation that the states had considerable latitude in this regard, that they could have, if they wished, banked the money, which is to say that they could have managed their actual expenditures if no ARRA monies had been forthcoming. Our conclusion is that even if they could have continued to spend, they wouldn't have; ARRA grants had a considerable impact on spending. But the stipulation must be understood as provisional: in the next section we argue, on the basis of the testimony of state budget officers, that most states could not have maintained their actual spending without the ARRA.

The cross-sectional tale is swiftly told. When we control for differences in financial solvency of the various states, ARRA grants explain a surprising amount of the cross-state variation in the changes in spending and the variation in the amounts added to state bank accounts between FY2009 and FY2010. The R^2's are of the order of 0.60 for the expenditure equations and 0.80 for the balance-sheet equations. The regressions suggest that some 2/3 of ARRA monies were spent by the states, the remaining 1/3 going to shore up state finances. Moreover, the results are reassuring (or surprising, depending on one's prior beliefs) in that the coefficients on the ARRA grants in the two equations add up to \$0.97 or \$0.99 in expenditure per dollar received, depending on the specification of expenditures and financial assets. In contrast with the procedure followed by Cogan and Taylor, this result is *not* because of a constraint that forces the coefficient on the net change in financial assets to unity, as in Tables 6 and 7, but the outcome of independent estimation of the determinants of expenditures and changes in financial balances. Finally, the results for expenditures are remarkably close to the numbers Christina Romer and Jared Bernstein employed in their prospective evaluation of the

ARRA.[16] However, our analysis gives little support for their idea that the ARRA would have a large effect on non-Federal revenues—that is, that the ARRA would dissuade states from raising taxes. Compared to the Romer-Bernstein assumptions, a much larger portion of the ARRA appears to have gone into increasing net financial assets *à la* Cogan and Taylor. The details are in Table 11.

Table 11. Regressions of Year-on-Year Changes in Expenditures, Revenues, and Assets on ARRA Grants, FY2010

	1	2	3	4	5	6
	Eqn 9	Eqn 10	Eqn 9	Eqn 10	Eqn 11	Eqn 9
Dependent Variables	Change in Total Expenditures	Change in Net Assets	Change in Current Expenditures	Change in Net Short Term Assets	Change in Non-Federal Revenues	Change in Total Expenditures
Total ARRA Grants Received in FY2010	0.656	0.337			-0.124	0.659
	0.071	0.147			0.053	0.072
Current ARRA Grants Expended			0.686	0.287		
			0.080	0.167		
Change in Non-Federal Revenues						0.094
						0.148
Net Financial Assets, Beginning of FY2010	0.00641	0.087			-0.025	.00737
	0.00320	0.007			0.009	.00356
Change in Net Financial Assets Between Beginning of FY2009 and Beginning of FY2010	0.236	0.429			-0.074	0.232
	0.044	0.090			0.041	0.044
Net Short Term Financial Assets, Beginning of FY2010				0.079		
				0.006		
Change in Net Short Term Financial Assets Between Beginning of FY2009 and Beginning of FY2010			0.180	0.462		
			0.033	0.096		
R^2	0.656	0.831	0.617	0.836	0.231	0.659
Adj R^2	0.634	0.820	0.601	0.826	0.175	0.630
Coverage	All States	All States	All States	All States	Non-Fossil-Fuel States	All States
N	50	50	50	50	44	50

[16] "For transfers to the states, we assumed that 60% is used to prevent spending reductions, 30% is used to avoid tax increases, and the remainder is used to reduce the amount that states dip into rainy day funds." (Romer and Bernstein, 2009, p 13)

Before turning to the analysis, we need to say a few words about the data. First, in contrast with the time-series analysis, the data here are restricted to state governments. There are several reasons for this. The most important is that the Census Bureau has not yet released state-by-state data that includes local governments beyond FY2009, and the ARRA had not disbursed much money when FY2009 ended (June 30, 2009 for all but four states). By contrast, the Census Bureau has published comprehensive data on state finances through FY2010. And the data are of better quality for the states than for the consolidated accounts of state and local governments; state government data are assembled from a survey of state governments and are not subject to sampling error, whereas local government data is collected through a sampling procedure. Another reason for focusing on the states is that the bulk of the ARRA monies paid out as grants to government entities, plus contracts and loans to non-government entities, in fact went to the states. Substantial amounts were in turn transferred to localities, as well as to higher educational institutions and other non-profits, by the states, but for reasons we have already considered, we regard these transfers as essentially equivalent to purchases of goods and services.[17]

For all the information on the recovery.gov website, no breakdown of ARRA grants is provided between states, localities, universities and other non-profits, and businesses. For the portion of grants covered by the recipient reporting requirement (Section 1512 of the ARRA), we separated the state grants by using a set of keywords like “department,” “education,” “executive office,” “human services.” For the programs not subject to Section 1512 reporting, the largest of which was Medicaid, we used the figures of the relevant Federal departments. Because the quarterly listing of recipient reports lumped together disbursements through September 30, 2009, we also relied on Federal agency reports of the Department of Education and the Department of Transportation to separate grants received by the states during FY2009 from grants received during FY2010.

[17] Although the supplement published by the Bureau of Economic Analysis on the impact of the ARRA only provides aggregate data for state and local governments, NIPA data breaks down Federal grants between states and localities. These data show only a very modest increase in *total* Federal grants to localities over the period of the ARRA. It follows that ARRA grants to the localities could not have been very large. This is confirmed by our analysis of the detailed ARRA data available on the recovery.gov website. Our calculation is that of the total grants, contracts, and loans reported through the end of calendar 2011 (plus Medicaid), 85 percent went to state governments, the rest going to private nonprofit entities (like universities), private businesses, as well as localities.

The general idea of the regressions reported in Table 11 was to use variations in spending among the several states to test the impact of the ARRA. If the ARRA had an impact, it should show up in greater expenditures by states receiving more ARRA money. If the ARRA had an impact on state balance sheets, it should show up in larger additions to net financial assets for states receiving more ARRA money. We also tested whether or not greater ARRA funding was associated with smaller changes in taxes and charges.

The general structure of the estimating equations is

$$\Delta O = a_0 + a_1 A + a_2 N_{-1} + a_3 \Delta N_{-1} + \varepsilon \qquad (9)$$

$$\Delta N = b_0 + b_1 A + b_2 N_{-1} + b_3 \Delta N_{-1} + \mu \qquad (10)$$

$$\Delta T = c_0 + c_1 A + c_2 N_{-1} + c_3 \Delta N_{-1} + \xi \qquad (11)$$

where

ΔO = Change in expenditure per capita, FY 2010 – FY2009 (expenditure = the sum of purchases of goods and services and transfer payments)
A = ARRA grants to states per capita of state population as of April, 2010
N_{-1} = Net financial assets per capita, beginning of FY2010
ΔN_{-1} = Change in net financial assets per capita during FY2009, $N_{-1} – N_{-2}$
ΔN= Change in net financial assets per capita during FY2010, $N – N_{-1}$
ΔT = Change in non-Federal revenues, FY2010 – FY2009 (taxes, charges, and miscellaneous revenues)

The null hypothesis, deriving from the work of Cogan and Taylor, is

$$H_0: a_1 = 0; b_1 = 1, c_1 = 0$$

The expenditure regressions, columns 1 and 3, imply that for each dollar of ARRA funding, between \$0.66 and \$0.69 was spent, depending on the inclusiveness of the concept of expenditure and the associated measure of ARRA grants, with the estimated value of the coefficient approximately 8 standard errors away from its null hypothesis value of 0. (The lower figure in column 1 is associated with the more inclusive measures, which include spending on capital account in expenditures and include spending of ARRA monies on infrastructure in the measure of ARRA grants. In arriving at the higher figure in column3, expenditure is limited to current account spending and ARRA grants are correspondingly limited to current expenditure.) The corresponding regressions of changes in net financial assets imply that between \$0.34 and \$0.29 of each dollar of ARRA money was added to the state's balance sheet. These estimates are approximately 4 standard errors from the null hypothesis value, $b_1 = 1$. (For the more inclusive definition of expenditure and ARRA funding, the measure of net financial assets is total assets less total liabilities. When expenditure and ARRA funding are limited to current account, the measure of net financial assets is limited to short-term financial assets.)[18]

To test the proposition that the ARRA affected taxation as well as expenditure and saving, we ran the regression reported in column 5. We limited the regression to the 44 "non-fossil-fuel" states—excluding Alaska, Louisiana, North Dakota, Texas, West Virginia, and Wyoming. It seemed to us inappropriate to include these six states, for which energy production provides them with a very different tax base, with the rest of the country. For instance, in FY2010 the 44 non-fossil-fuel states obtained almost 50 percent of their non-Federal revenues from a combination of sales taxes and individual and corporate income taxes, whereas the six energy states relied on this combination for only 30 percent of their non-Federal revenues. Fossil-fuel states also differ from the rest of the country with respect to their balance sheets: at the beginning of FY2010 the combined assets of the six were 266 percent of liabilities; for the other 44 states assets were only 87 percent of liabilities. (Data from the Census Bureau, 2012; the combined balance sheet of all the states masks the great disparities: for all 50 states together assets were 99.8 percent of liabilities.)[19]

For the 44 non-fossil-fuel states, the ARRA appears to have diminished the need for additional revenue from non-Federal sources. The coefficient on the ARRA is −0.124, and the standard error allows us to reject the null hypothesis of $c_1 = 0$. But the limitation of coverage of this regression makes it difficult to integrate the result with the estimates of a_1 and b_1. Observe

---

[18] Table 11 omits variables for which preliminary regressions, not reported here, gave insignificant results: constant terms in all regressions were omitted for this reason, as was short term financial assets in column 3.

[19] Why don't we, by the same logic, limit the expenditure and saving regressions to the 44 non-fossil-fuel states? We ran regressions on the 44-state subsample (not reported here), but the results for the total expenditure regression differed very little from the same regression on the full data, while the coefficient on current ARRA funds was substantially reduced in the regression we ran on the restricted set of data. In the saving regression run with total expenditures and total ARRA grants, the coefficient on ARRA funds is no longer precise enough to shed much light on the null hypothesis, $b_1 = 0$. In the regression relating non-Federal revenues to ARRA grants, the sign of the coefficient on total ARRA funds has the wrong sign and is insignificant, whereas the in the current-account version of the regression (reported in column 6 of Table 11) the coefficient on the ARRA variable has the expected negative sign, and although small in magnitude is significantly different from zero. The lack of variation among the non-fossil-fuel states is presumably the source of the loss of precision.

that the coefficients on the two control variables are also of the sign we would anticipate in a world in which state budget policy is partly driven by balance-sheet considerations.

The regression reported in column 6 is intended to test the hypothesis that non-Federal revenue influenced expenditure decisions in FY2010. The coefficient is of the right sign, but it is numerically small, 0.09, as compared with coefficients on ARRA money of 0.66 for total grants and 0.69 for current grants. Moreover the coefficient on changes in non-Federal revenue differs insignificantly from 0.

Finally, Table 12 reports the mixed results of including political party as a determinant of budget behavior. These regressions add a dummy variable for the states whose governors were Republicans at the beginning of FY2010.

Table 12. Regressions of Year-on-Year Changes in Expenditures, Revenues, and Assets on ARRA Grants and Political Party, FY2010				
	1	2	3	4
	Eqn 9	Eqn 10	Eqn 9	Eqn 10
Dependent Variables	Change in Total Expenditures	Change in Net Assets	Change in Current Expenditures	Change in Net Short Term Assets
Total ARRA Grants Received in FY2010	0.765	0.396		
	0.088	0.189		
Current ARRA Grants Expended			0.770	0.379
			0.102	0.213
Change in Non-Federal Revenues				
Net Financial Assets, Beginning of FY2010	.00697	0.088		
	.00312	0.007		
Change in Net Financial Assets Between Beginning of FY2009 and Beginning of FY2010	0.238	0.430		
	0.042	0.091		
Net Short Term Financial Assets, Beginning of FY2010				0.079
				0.006
Change in Net Short Term Financial Assets Between Beginning of FY2009 and Beginning of FY2010			0.181	0.472
			0.033	0.098
Republican Governor	-80.652	-43.554	-49.604	-55.635
	38.879	85.763	37.455	79.423
R^2	0.684	0.832	0.630	0.838
Adj R^2	0.657	0.817	0.607	0.824
Coverage	All States	All States	All States	All States
N	50	50	50	50

The regression reported in column 1, for total expenditures, suggests that political partisanship played a role in how much of the ARRA monies were spent. A Republican in the governor's chair reduced spending by \$81 per capita! (This is a relatively large amount relative to the average change in per-capita spending across the 50 states, \$140, the change ranged from a low of −\$269 in Alaska to \$921 in North Dakota, both states presumably heavily influenced by energy revenues and the effects of these revenues on their balance sheets.) But the effect doesn't hold up for the regressions on the amount of ARRA money saved, reported in columns 2 and 4. Here the expected positive effect is belied by the negative coefficient, and in any case the standard errors are too high to put very much credence in these estimates. Similarly, although in the current account specification of the expenditure equation the coefficient of the party-affiliation variable has the right sign, it too is characterized by a high standard error.

There remains the possibility that the causality in the equations represented in columns 1 and 3 run from higher spending to higher grants, rather than the other way around. Reverse causality would leave the interpretation of the impact of the ARRA open: it is still possible that the ARRA stimulated state expenditure—the states spent more in the expectation of reimbursement than they would have if there had been no ARRA [20]—but it is also possible that the states would have spent the same amounts with or without ARRA. However, this possibility of reverse causality can be rejected for two reasons. First, there is no parallel argument for reverse causality in the relationship between ARRA grants and state balance sheets, so this cannot be the explanation for the statistical rejection of the second part of the null hypothesis, namely $b_1 = 1$. And if causality runs from the ARRA to state balance sheets for 1/3 of the grant, it ought to run in the same direction with respect to expenditure, especially since the two coefficients together get us to within spitting distance of $1.00. The second reason is that regressing the error term against the ARRA variable reveals no correlation. The two pairs of regressions in Table 13, the first with and the second without a constant term, indicate no correlation between the error term and the level of the ARRA grant: the constant term is insignificantly different from 0 in the first regression and the correlation between the error term and the ARRA variable is 0 when the constant term is forced to 0, in the second regression.

[20] This is how the Medicaid program works, Federal reimbursement depending in part on how much states commit to the program.

Table 13. Regressions of Residuals from Eqn 9 on ARRA Grants, FY2010				
	1	2	3	4
Dependent Variables	Residual	Residual	Residual	Residual
Constant	-33.914		26.105	
	91.948		78.983	
ARRA Grants Received	0.104	0.000		
	0.290	0.063		
ARRA Current Expenditure			-0.098	0.000
			0.304	0.072
R^2	0.003	0.000	0.002	0.000
Coverage	All States	All States	All States	All States
N	50	50	50	50

At this point, it seems fair to conclude that the econometric evidence does not support Cogan and Taylor. Expenditure smoothing is either a bright idea whose time has not yet come, or a misplaced faith in the rationality and liquidity of economic agents.

4. What State Budget Officers Say About the ARRA

There are limits to the inferences one can make from econometric results. The interpretation of regression coefficients, standard errors, and the rest requires a framing theory—including, crucially, a set of counterfactual assumptions. Cogan and Taylor interpret their results against the counterfactual assumption that states behave like the consumption-smoothing individuals theorized by Friedman and Modigliani in the permanent-income and life-cycle hypotheses. Specifically, they assume that in the absence of ARRA "the states would have held government purchases at the levels actually observed during the recession and would have instead not increased net lending as they did during this period" (Cogan and Taylor 2011b, p. 12; this sentence is omitted from subsequent versions of the paper, including the published version, 2012). It is against this assumption that the observed lack of significant increase in aggregate state government purchases from FY2009 to FY2011 is interpreted as evidence that the ARRA was a failure.

In the absence of this assumption the econometric results could be interpreted very differently. For example, the *prima facie* equally plausible assumption that states would have sharply cut spending in response to the recession, below the levels actually observed, would support precisely the opposite interpretation.

We have already seen that the time-series evidence for Cogan and Taylor's interpretation of the data is wanting in that whether or not expenditure is driven by revenues (including ARRA revenues), time-series regressions would indicate a large and significant coefficient on lagged outlays. And if indeed the correlation between lagged and current outlays is spurious, the coefficient on revenues would be biased downward. But it is one thing to argue that the time-series data do not support Cogan and Taylor and another to argue that these data reject their hypothesis. They do not. And examination of the aggregate data during previous downturns provides evidence for and against their view. For this reason, the previous section turned to cross-sectional evidence. In our judgment the cross-sectional data provide strong evidence against Cogan and Taylor. Specifically, cross-section regressions indicate that 2/3 of the ARRA money was spent by the states. But we would be the last to claim that our analysis is conclusive.

We therefore supplement these regressions with a direct examination of the plausibility of their counterfactual assumption: a set of open-ended interviews with state budget officers. We sent a questionnaire to all fifty state budget officers that gave the respondents substantial latitude in their answers. The questions were framed to provide a foundation for conversation without being so restrictive that they would prevent us from learning things about state budgeting practice that we had not anticipated.

We recognize the unorthodox aspects of this approach. Economists generally resist asking agents for information about why they do what they do or what they would have done if the circumstances had been different.[21] Often, there is good reason for this reluctance: there are too many agents, it is hard to get a representative sample, and agents may have trouble reconstructing the circumstances of their decisions well enough to answer, especially when the questions involve a counterfactual. Fortunately, none of these reasons apply to the case at hand. There are only 50 states, and state budget officers are a highly professional group of men and

[21] There are a few notable exceptions. For example, Henderson (1938) and Meade and Andrews' (1938) use of interviews with businessmen to explore the impact of the interest rate in the determination of investment; and Blinder (1998) and Bewley's (1999) discussions with relevant economic actors to explore the reasons behind the stickiness of prices and wages, respectively.

women. *A priori*, then, it seemed sensible to us to ask these officers what they would have done had there been no ARRA funds to offset lost revenues and increased demands for expenditure that were the twin results of the Great Recession. From the information we gathered, we conclude that Cogan and Taylor's assumption that in the absence of ARRA states could have and would have increased net borrowing to fund spending at roughly the levels observed *with* ARRA is highly implausible, and that it is much more plausible that the great majority of the states would have cut spending significantly without ARRA. In the remainder of this section, we present the evidence in support of this conclusion.

Study design

Our goal was to elicit responses from all fifty state budget directors to eight questions designed to allow us to assess the plausibility of Cogan and Taylor's counterfactual assumption. The questions were as follows:

1. What would have been the consequences for current and capital spending had no ARRA money come to [your state]?
2. Again, assume that no ARRA money had come to [your state]. In this case, would your capital budgeting process have required you to reduce capital expenditures in response to worsening economic conditions?
3. Outside of the general fund (and stabilization funds) were there other options for funding current budget deficits that might have arisen without ARRA? (*For example: special funds from other public of quasi-public agencies not included in the general fund, but that can be drawn on by the state? Revenue anticipation notes or similar instruments?*)
4. Is it possible for [your state] to borrow to finance operating-budget deficits?
5. In your experience, did the maintenance of effort provisions (MOE) attached to ARRA funding significantly restrict [your state]'s flexibility regarding how to use the funding?
6. In [your state], is there any flexibility with regard to classifying expenses as "current" or "capital"?
7. In [your state], is any portion of the capital budget typically funded from the operating budget (i.e., using current revenues as opposed to bonds)?

8. With regard to ARRA funding for capital projects: to the extent you received such funding, did it fund new incremental capital spending, or did it just act as a replacement funding source?

As explained above, we did not intend the questions to be, in themselves, comprehensive and complete. Rather, we intended them to be a foundation for a less structured provision of information. We specifically wanted to avoid biasing the answers by rigidly steering the budget directors to respond only to those issues that *we* felt were relevant and important and thereby closing off issues that we had not anticipated. At the same time, we wanted the questions to articulate the specific issues relevant to assessing the Cogan-Taylor counterfactual. We chose the questions with the aim of balancing these two imperatives. In order to make participation as convenient as possible we offered the budget officials the option of answering the questions by e-mail or through a brief phone interview.

Composition of respondents

Of the 50 state budget directors we contacted, we received responses or had phone interviews with 29. Obviously, our aim was to collect information from all of the states and we made efforts over a 5 month period to collect a comprehensive set of responses. Despite these efforts, however, we received no response from 21 states. Nonetheless, we feel that our group of respondents is large and comprehensive enough and similar enough to the non-response set in many important demographic aspects to give us some confidence that the responses are not tainted with selection bias.[22] Table 14 contains demographic and economic summary statistics of the two groups.

---

[22] Of course, due to the qualitative and relatively open-ended nature of the information being gathered, we cannot formally quantify the extent of bias or confidence in our conclusion.

Table 14. Demographic and Economic Comparison of Responding and Non-Responding Groups

	Responding States	Non-Responding States	All States
GDP (Bil of 2005 $)	$8,290	$4,716	$13,006
Pop. (Mil, July 2009)	187.1	119.3	306.4
GDP/Capita	$44,307	$39,537	$42,447
Total Expend. ($Bil)	$998	$559	$1,557
Repub Gov	48%	71%	58%

Findings and interpretation

The main thrust of our first question was simply to ask budget officers directly if it struck them as plausible for their state that they could have maintained expenditures in the absence of the ARRA. There was virtual unanimity among the respondents that this was *not* plausible. Since the evidence supporting the implausibility of business as usual in the face of the Great Recession is slightly different for operating and capital expenditures, we will discuss the two separately, beginning with operating expenditures.

Of all of the respondents, only those states with significant fossil-fuel related revenues indicated that it either might have been possible or definitely would have been possible for them to maintain operating expenditures in the absence of ARRA. Alaska and North Dakota's oil and gas revenues, respectively, shielded them more or less entirely from the budgetary woes of the recession. The response from North Dakota, where the unemployment rate never went above 4.2 percent and was under 4 percent for most of the period in question, was essentially "Recession? What recession?" West Virginia was also shielded, though not quite as thoroughly due to very high Medicaid costs that they might not have been able to cover without ARRA's enhancement of the Federal portion of Medicaid costs (that is, the Federal Medical Assistance Percentage or FMAP). Wyoming avoided the worst of the recession both through fossil-fuel related revenues and two rounds of expenditure cuts in the lead-up to the recession that remained in place throughout. These states, however, were the exceptions. All of the other respondents indicated that it would not have been possible to maintain expenditures at the observed levels in the absence of ARRA without additional revenue-raising measures (increased taxes and/or fees).

Michigan's State Office of the Budget, for example, reported that "[h]ad no ARRA funding come to Michigan, general fund reductions of approximately 18% would have been required each fiscal year and would have been in addition to measures taken to close a $1.4 billion budget gap for fiscal 2009, and $1.8 billion in general fund reductions enacted for fiscal 2010." Moreover, many of the respondents commented that it was likely that the balance of the adjustment to lower revenues would likely have been weighted heavily toward spending cuts rather than tax or fee increases due to political considerations.

The sentiment that lower operating expenditures would have been necessary without ARRA was not sensitive to political party—it was voiced equally by those states with Democratic and those with Republican governors. There was, however, some difference along political lines with respect to the attitude toward the maintenance of spending that was enabled by ARRA. Several officials from Republican states told us that while their states would likely have enacted more spending cuts in the absence of ARRA, this would have been a positive rather than a negative for economic health.[23] We heard this comment both with respect to spending in general, and specifically with respect to Medicaid and education—two areas where ARRA money came with maintenance of effort ("MOE") provisions. In general, the theme of these comments was that ARRA allowed the state government to put off dealing with budgetary problems, some of which were structural and would still have to be dealt with once the ARRA funds dried up. Many of the budget officials commented that they were wary of creating a "fiscal cliff" by using ARRA money to continue to fund programs at levels that would likely be unsustainable post-ARRA.

The responses of the budget officials regarding operating expenditures takes into account the possible impact of budget stabilization funds (BSFs)—i.e. that even with the aid of internal reserves it would not have been possible to maintain expenditures in the absence of ARRA. Again, with the exception of the fossil-fuel states, all of the respondents commented that their

[23] This sentiment was expressed to us by state budget officers whose current administration is Republican—in particular, those from Ohio, Wyoming and Kansas—or whose state was under a Republican administration during the years in question—in particular, Minnesota.

BSFs would not have been sufficient to have undertaken spending at the observed levels.[24] During fiscal 2009, for example, Minnesota drew its budget reserves down to zero and was projecting revenue shortfalls several years in to the future. Similarly, Arkansas, which did not create a budget stabilization fund until 2010, faced significant revenue shortfalls in FY 2010 and would not have been able to support the operating budget actually executed in 2010 without ARRA. In this connection, it is important to note that the vast majority of the respondents—including those from the fossil fuel states—indicated that they would have made significant efforts to avoid drawing their BSFs down to zero. This was important to them for two reasons. First, these funds are an important bulwark against all kinds of fiscal emergencies, and not just recessions. Iowa, for example, faced significant unexpected expenses from a major flooding episode in the summer of 2008. Second, as a former Massachusetts state budget officer indicated, the level of these reserve funds affects a state's credit rating. This adds an additional potential cost to drawing them down too far.

While drawdowns from BSFs are the most obvious form of covering revenue shortfalls, there are other possibilities as well, at least in principle. For example, states may have special funds (for example, from lottery revenues or transportation-related fees) that are a part neither of the general fund nor the BSF that could in principle be tapped to fill general revenue shortfalls. These would simply be another form of "reserve drawdown" and would therefore qualify as additional net borrowing by Cogan and Taylor's definition. The responses to Question 3 provided direct evidence about this possibility. Nineteen of the respondents indicated either that no such funds would have been available to cover general revenue shortfalls or that such funds could have been tapped but the amounts would have been insignificant. Five states indicated that such funds exist, are substantial and can be tapped, but that even with these contributions the revenue shortfall would have been too great to meet without additional measures (absent ARRA). A Maryland official commented that although "reprioritizing special funds is a significant tool in budget balancing…it would not have been sufficient to prevent significant reductions to key state services in the absence of ARRA funds." Three additional states indicated that such funds were already exhausted or being used to the greatest extent possible during the

[24] The one exception to this was Oklahoma, which has significant fossil-fuel related revenues, but indicated that they drew their reserve funds down steadily to zero over the course of fiscal 2010 and 2011.

period of ARRA funding. A Connecticut official commented that during the recession, special funds “were significant in offsetting the State’s large shortfalls”, but that “by 2011 all fund sweeps had been exhausted. [And] ARRA filled part of the gap.” And two states, both fossil fuel states, indicated that the point was moot because they would not have needed to explore such possibilities.

Another possibility open to states, in principle, to maintain operating costs through increased net borrowing would be to shift operating costs onto the capital budget. Question 6 asked the budget officials if such a strategy was open to them. With the exception of three states—Hawaii, Kansas and Utah—all respondents indicated that there is very little such flexibility and, that to the limited extent such shifts could be made, their impact would not be significant.[25] In the case of the three exceptions, officials from Kansas and Hawaii indicated that there was some flexibility to make such reclassifications during the recessionary period but did not indicate how significant the flexibility was or the extent to which it was used. And the Utah respondent indicated that some building projects that had been on the operating budget were shifted to the capital budget.

These responses, then, directly undermine the plausibility of the Cogan-Taylor counterfactual with respect to operating expenditures. The responses show that the possible borrowing sources—internal reserve funds and capital market borrowing—either would not have been sufficient, ex-ARRA, to fund operating expenditures at the observed level or were not available for that purpose.

For purposes of assessing the plausibility of the Cogan-Taylor counterfactual, the effect of ARRA on capital expenditures is more complicated than its effect on operating expenditures. The primary reason for this is that direct capital grants constituted a relatively small portion of total ARRA outlays to the states, and that the other portion of the funds—that is, the vast majority of ARRA outlays to states—affected capital spending in less direct and more complex ways. It may be useful to begin by working out what would count as evidence for and against the Cogan-Taylor counterfactual before turning to the responses.

[25] Many respondents indicated that some operating expenses directly related to capital projects (for example, the salaries of personnel dedicated to the project in question) are routinely included in the cost of the capital project.

In order for their counterfactual to be plausible with respect to capital spending, it must be the case that the states would have had the wherewithal to fund capital expenditures at the observed levels in the absence of ARRA. For the vast majority of the states capital expenditure is funded with debt, mostly in the form of bonds (for example, general obligation bonds and revenue bonds). For those that fund capital expenditures largely or entirely from the operating budget, our assessment of the effect of ARRA on operating expenditures carries over to capital expenditures: capital expenditures funded out of the operating budget are equivalent to operating expenditures for our purposes.[26] In light of this, our assessment of the plausibility of the Cogan-Taylor counterfactual hinges on whether or not the debt-financing states could have borrowed enough, absent ARRA, to support the capital expenditure actually observed.

To determine whether or not this is the case, we need to consider all of the various paths through which the ARRA might have affected both borrowing capacity and observed capital expenditure. In all—again, for those states that fund their capital expenditures with debt—there are three: (a) ARRA funds designated specifically as capital grants could have been used directly (i.e. without borrowing) to fund capital projects that otherwise would not have been funded; (b) states could have taken advantage of ARRA's "Build America Bonds" program—which provided a partial subsidy for states' interest payments on newly issued taxable bonds eligible for the program—to access the capital markets to a greater extent than otherwise might have been possible; and (c) the increased revenue from ARRA could have increased the state's borrowing capacity above what it would have been without those revenues.

With respect to direct capital grants, virtually all of our respondents indicated that ARRA had allowed them to undertake incremental capital spending—either in the form of new projects or the acceleration of existing planned projects. In Florida, for example, ARRA capital grants were used by the Florida Department of Transportation to fund "more complex projects which would result in higher job creation." They further indicated that this funding "resulted in projects

[26] These states are Alaska, North Dakota, West Virginia, Wyoming and Arkansas. The first four are fossil fuel states for whom it is plausible to suggest that they would have been able to undertake the observed level of capital expenditure absent ARRA (for the same reason this was deemed plausible for operating expenses). The Arkansas respondent, however, indicated that ARRA funding allowed them to undertake certain critical infrastructure projects that would not have been possible otherwise. In addition, several other states—Delaware, Florida, Iowa and Rhode Island—allocate some amount of surplus general revenue funds, when available, to capital projects funds various kinds. During the recession, these funds were largely exhausted.

in addition to state funded projects planned for expenditure during this period." Hawaii and Ohio were partial exceptions, with the respondents of these states indicating that ARRA funds were at least in part used as replacement funding, though the precise extent of replacement was unclear.

With respect to the impact of the Build America Bonds (BAB) program, the responses did not provide enough information to form a clear conclusion. Only four of the states explicitly mentioned the BAB program as having had a significant impact. Of these, two—California and Colorado—indicated that the program had allowed them to undertake more capital spending than would have been possible in the absence of the program. BAB was especially important to California, which was having difficulty accessing the credit markets through standard channels. And an official from the Colorado budget office indicated that, due to the BAB program, the Colorado Bridge Enterprise "actually issued $300 million in bonds that it otherwise may not have issued." With respect to the other two states, Rhode Island indicated that Build America Bonds were used for refinancing purposes, and Ohio indicated that the BAB program was probably not used for any capital spending that would not have been undertaken in any event.

With respect to the effect of ARRA on states' borrowing capacity, we found that ARRA generally did not affect borrowing capacity. For most of the states, annual capital borrowing is capped by statute, with the cap generally being related in some way to projected revenues. [27] In all of our responding states where this is the case, ARRA revenues were not included in "projected revenues" for the purposes of the debt limit calculation. So the amount these states were willing and able to borrow each year would not have been affected by ARRA grants.

Putting all of these pieces of the story together, it is reasonable to conclude that—at least for the responding states that fund capital borrowing with debt—they could not have undertaken capital spending at the observed levels in the absence of ARRA. Since ARRA led to incremental capital spending in almost all of these states, actual capital expenditures were greater than what they would have been in the absence of ARRA. Since states' ability to increase capital borrowing was limited by statute—with no responding states indicating that they were borrowing at a rate

[27] In Massachusetts, for example, the debt cap was set by statute in 1990 to be $6.8 billion, and to grow by 5% each year. The same legislation limited total annual debt service (interest and principal) on state general obligation debt to no more than 10% of budgeted appropriations (M.G.L. Part I, Title III, Chapter 29, Sections 60A, B). See Commonwealth of Massachusetts, "Debt Affordability Analysis," Report of the Executive Office of Finance and Administration (http://www.mass.gov/bb/cap/fy2009/dnld/fy09capappendixama.pdf) for an analysis for FY 2009-2012 that utilizes these guidelines.

below that limit—it would not have been possible for them simply to fund such incremental expenditure with additional capital market borrowing.

The evidence gleaned from our questionnaire to state budget officials, then, supports the conclusion that, contrary to the claims of Cogan and Taylor, it is *not* plausible to claim that in the absence of ARRA states would have undertaken expenditures at the level actually observed and would have funded this with additional net borrowing. The evidence indicates that states not only *would not* have done so, but that in almost all cases they *could not* have done so. The borrowing sources available to states are the (financial) capital markets and their internal reserves, which include budget stabilization funds (including rainy day funds) and any other special funds that are available for filling shortfalls. The responses to our questionnaire indicate that, aside from the fossil fuel states, internal funds would have been insufficient to support operating expenditure at the level actually observed and that maintaining that expenditure by shifting it to the capital budget was not an option available to a large enough extent to have made a difference. On the capital expenditure side, the responses indicated that (in the absence of changes to statutes) states are very limited in their ability to increase capital borrowing during recessions as their capital borrowing limit is tied to projected revenues.[28] In light of this, it is not plausible to interpret the observed data on expenditures and net borrowing as evidence that ARRA grants to states failed to stimulate additional spending relative to the state of the world without those grants.

5. Conclusions

Did the stimulus work? Our short answer is yes. But to answer the question of whether or not the ARRA stimulus—or any fiscal stimulus for that matter—"worked" one needs to be clear about the assumptions that provide the framework within which the results are interpreted. Based on our analysis, the proper counterfactual assumption about state government expenditure during the recession is that in the absence of ARRA states would have been unable to maintain expenditures at (or close to) pre-recession levels. The proper measure of success, therefore, is not an observed rise in the expenditure trend, but instead evidence that ARRA funds were used for incremental spending relative to that counterfactual. On that basis, our analyses support the

[28] Again, the fossil fuel states are an exception to this, as they generally fund their capital expenditures from the operating budget. The question of whether they can engage in counter-cyclical capital borrowing, then, is not pertinent.

conclusion that the stimulus worked, and that the "rational" expenditure smoothing arguments to the contrary are invalid. The evidence suggests that state government expenditure was significantly increased by ARRA relative to what would have been possible without it. For the economy more broadly, the evidence suggests over the period from mid-2009 to mid-2011 it added some 2 percent to GDP. If the Obama Administration can be faulted, it would be for failing to appreciate the gravity of the situation it inherited in January 2009, for lacking the courage or foresight to ask for more stimulus over a longer time period, for failing to argue forcefully enough that more of the stimulus should be directed to lower income beneficiaries who would have been more likely to spend than to save, or for all three.

Detailed examination of the evidence with respect to grants to the states reinforces our admittedly more casual evaluation of the stimulus as a whole: contrary to the conclusions of Taylor (2011) and Cogan and Taylor (2012), both econometric analysis of cross-sectional state data and interviews with state budget officers suggest that the ARRA allowed the states to maintain spending programs that would have been drastically cut if the stimulus had not been enacted. A portion of the ARRA monies did go to shore up state balance sheets—as indeed was the intention of the ARRA legislation—but far less than Cogan and Taylor contend. Our estimate, based on cross-sectional data, is that during FY2010 approximately 1/3 of grants to states made under the ARRA were added to their balance sheets, whereas 2/3 were spent.

Beyond the stimulus, an important lesson of this paper is the need for methodological pluralism. Aggregate time series is one source of evidence, but as we have seen, it is unnecessarily limiting to focus exclusively on this particular evidence. Bringing cross-sectional and interview evidence to bear adds considerably to our understanding of the impact of the ARRA on state finances.

But the chief methodological lesson is the absolute necessity of grounding empirical assessments firmly in empirical reality, subjecting one's assumptions to rigorous scrutiny *by whatever investigative means are required*. This is necessary to avoid misinterpreting econometric results—even those arrived at through impeccable econometric analysis. Cogan and Taylor hypothesize that state governments and individual agents engage in expenditure smoothing. This is a reasonable hypothesis for many purposes, one with at least two Nobel Prizes on its side. But the *scope* of its legitimacy is circumscribed by the assumptions on which the underlying theory is based. Before one can legitimately deploy it as a counterfactual assumption, one must know enough about the empirical reality of the target population to ensure that it is really plausible.

The importance of this methodological caveat is especially clear in the case at hand, where the discourse over the success or failure of ARRA has been rife with self-fulfilling analyses on both sides of the debate. Taylor rightly argues that most of the post-hoc vindication of the ARRA could have been—and in fact was—written before one dime of ARRA monies had been spent. But it is equally true that this a case of the pot calling the kettle black: whatever the truth of the Cogan-Taylor hypothesis, their methodology guaranteed that the data on state and local governments would "confirm" expenditure smoothing by displaying a high correlation between current and lagged expenditure.

Although all empirical analysis presupposes a theoretical framework, not all frameworks are created equal with respect to how much room there is for the empirical results to contradict a preferred hypothesis. Our analyses of the state government channel of the ARRA were designed to minimize the extent to which the assumptions drove the results. In contrast with the Cogan-Taylor assumptions that guaranteed the appearance of expenditure smoothing, our cross-sectional regressions left the answer open: the regressions might have turned out very differently. Certainly there was no *a priori* guarantee of favorable results—the high R-Squares and associated t-values of the coefficients in Table 11, and more particularly that the independently estimated coefficients on outlays and asset accumulation sum to unity. Of course, as with all econometric analyses, our results must also be interpreted in light of our assumptions. Responsible econometricians never tire of pointing out that correlation does not imply causality.

Given these limits to the efficacy of regression analysis, we sought to shed additional light on the question of causality by eliciting information directly from SBOs—the very agents who would have been the vehicle of cause and effect. And here too we made every attempt to frame our questions in a way that would have permitted answers on both sides. The open-ended questions we posed allowed SBOs to range freely in their answers. There was a uniformity of responses—but *not* unanimity—with respect to how the ARRA actually affected expenditures, even when the respondents obviously differed in their evaluation of the ARRA as a policy. So, while it is true that the framework of analysis affects the results, it is not the case that all frameworks are created equal. We would claim that our own framework is less restrictive, more open to alternative outcomes, and more sensitive to empirical reality than the framework invoked by the leading critics of the stimulus.

The final lesson is skepticism about the conventional distinction between positive and normative economics. Taylor is very much in the mainstream in believing that description can be separated from values, the first representing science the second ideology. But just as there are no facts without theory, there is no separate realm for description that does not embody values. Ideology ought not to be, as it is glossed in the *Cambridge Dictionary of Philosophy* (1999, 2nd Ed., p 406) "a disparaging term used to describe someone else's political views which one regards as unsound." Acting on ideology is not a failing or disease of the Other against which Taylor (or we for that matter) can claim immunity. Ideology is not the coin of the realm of true and false. We all operate on the basis of assumptions that cannot be proved or disproved, and ideology is the coin of the vast realm of what is beyond our powers to confirm or deny. This does not mean there is nothing to discuss, nothing to learn. To the contrary. We may seek to transcend ideology, but we will never do so until we admit that it is the necessary starting point of any serious discussion about policy.

References

Achen, C. (2001). Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables. Unpublished working paper. Department of Political Science and Institute for Social Research. University of Michigan.

Ando, A.K and Modigliani, F. (1963). The "life cycle" hypothesis of saving: Aggregate implications and test. *American Economic Review* 53(1): 55-86.

Backhouse, R. and Bateman, B. (2011). *Capitalist Revolutionary: John Maynard Keynes*. Cambridge, MA: Harvard University Press.

Barro, R.J. (1989). The Ricardian Approach to Budget Deficits. *Journal of Economic Perspectives* 3(2):37-54.

———. (2009). Voodoo Multipliers. *Economists' Voice*, vol. 6, issue 2, article 5.

Bewley, T.F. (1999). *Why Wages Don't Fall During a Recession.* Cambridge, MA: Harvard University Press.

Blinder, A.S. (1998). *Asking About Prices: A new approach to understanding price stickiness*. New York: Russell Sage Foundation.

Cochrane, J.H. (2009). Fiscal Stimulus, Fiscal Inflation, or Fiscal Fallacies? faculty.chicagobooth.edu/john.cochrane/research/papers/fiscal2.htm. Accessed 6/24/12.

———. (2010). Stimulus: Neither needed nor free. *Los Angeles Times*, online edition. http://articles.latimes.com/2010/nov/14/opinion/la-oe-economists-20101114-web.

Cogan, J.F. and Taylor, J.B. (2011a). Stimulus Has Been a Washington Job Killer. *Wall Street Journal*. October 4, p. A21.

———. (2011b). What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package. Working paper, January, 2011.

———. (2012). What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package. In *Government Policies and the Delayed Economic Recovery*, Lee. E. Ohanion, John B. Taylor and Ian J. Wright, eds. Stanford, CA: Hoover Institution Press.

Congressional Budget Office (2012). Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output from October 2011 Through December 2012: A CBO study (CBO Publication No. 4435). Washington, DC: U.S. Government Printing Office. Retrieved from http://www.cbo.gov/sites/default/files/cbofiles/attachments/02-22-ARRA.pdf.

Davis, O.A., Dempster, M. A. H. and Wildavsky, A. (1966). A Theory of the Budgetary Process. *The American Political Science Review* 60(3): 529-547

———. (1974). Towards a Predictive Theory of Government Expenditure: US Domestic Appropriations. *British Journal of Political Science* 4(4): 419-452.

Feldstein, M.S. (2009). Rethinking the Role of Fiscal Policy. *NBER Working Paper* #14684. http://www.nber.org/papers/w14684

Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.

Gramlich, E.H. (1978). State and Local Budgets the Day After it Rained: Why is the surplus so high? *Brookings Papers on Economic Activity* 1978(1): 191-216.

———. (1979). Macro Policy Responses to Price Shocks. *Brookings Papers on Economic Activity* 1979(1): 125-166.

Henderson, H. D. (1938). The Significance of the Rate of Interest. *Oxford Economic Papers* 1: 1-13.

Initiative on Global Markets (2012). IGM Forum: Economic Stimulus. University of Chicago, Booth School of Business. February 15. http://www.igmchicago.org/igm-economic-experts-panel/poll-results?SurveyID=SV_cw5O9LNJL1oz4Xi. Accessed 6/24/12.

Keele, L. and Kelly, N. (2006). Dynamic Models for Dynamic Theories: The ins and outs of lagged dependent variables. *Political Analysis* 14(2): 186-205.

Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money*. London: Macmillan.

Marglin, S.A. (2008). *The Dismal Science: How thinking like an economist undermines community*. Cambridge, MA: Harvard University Press.

Meade, J.E. and Andrews, P.W.S. (1938). Summary of replies to questions on effects of interest rates. *Oxford Economic Papers* 1: 14-31.

Modigliani, F. and Brumberg R.H. (1954). "Utility Analysis and the Consumption Function: An interpretation of cross-section data." In *Post-Keynesian Economics*, ed. Kenneth K. Kurihara, 388-436. New Brunswick, NJ: Rutgers University Press.

National Association of State Budget Officers (2012). *The Fiscal Survey of States*. Spring. Washington, DC: National Governors Association and the National Association of State Budget Officers.

National Public Radio (2011). Boosting Job Growth: It's a money matter. *Weekend Edition Sunday*. August 14. Podcast, available at http://www.npr.org/2011/08/14/139617127/how-can-washington-boost-job-growth.

Ramey, V. (2011). Can government purchases stimulate the economy. *Journal of Economic Literature* 49(3): 673-685.

Romer, C. and Bernstein, J. (2009). The Job Impact of the American Recovery and Reinvestment Plan. http://otrans.3cdn.net/45593e8ecbd339d074_ l3m6bt1te.pdf. Accessed 6/24/12.

Stack, C. (1975). *All Our Kin: Strategies for survival in a Black community*. New York: Harper & Row.

Stiglitz, J. (2010). Federal Spending is a Necessity. *Los Angeles Times*, online edition, http://articles.latimes.com/2010/nov/14/opinion/la-oe-economists-20101114-web.

Taylor, J.B. (2011a). An Empirical Analysis of the Revival of Fiscal Activism in the 2000s. *Journal of Economic Literature* 49(3): 686-702.

———. (2011b). The 2009 Stimulus Package: Two years later. Testimony before the Committee on Oversight and Government Reform, Subcommittee on Regulatory Affairs, Stimulus Oversight and Government Spending, U.S. House of Representatives. February 16.

http://media.hoover.org/sites/default/files/ documents/2009-Stimulus-two-years-later.pdf. Accessed 6/24/12.

Wildavsky, A. (1964). *The Politics of the Budgetary Process.* Boston: Little Brown.