Policy Analysis

Why in the World Are We All Keynesians Again?
The Flimsy Case for Stimulus Spending

by Andrew T. Young

Executive Summary

The U.S. federal government responded to the financial crisis and recession that began in 2007–08 with unprecedented fiscal stimulus. Passed in February of 2009, the American Recovery and Reinvestment Act (ARRA) came with a price tag of $831 billion. Yet the economy has not returned to a path of robust economic growth and unemployment has stubbornly remained high; only in September of 2012 did it dip (barely) below 8 percent. This has not stopped the Obama administration from pushing for further fiscal stimulus.

Whether or not fiscal spending stimulus is effective hinges critically on the size of the *spending multiplier*, which is dependent on several factors. In particular, if individuals anticipate the future tax liabilities associated with deficit spending and/or are "crowded out" by the deficit spending, then the multiplier is likely to be less than one; that is, each dollar of stimulus increases total spending in the economy by less than one dollar. Ultimately, whether the government spending multiplier is less than or greater than one is an empirical question.

While no one seriously debates the long-run costs associated with exploding public debt, the evidence suggesting significant short-run benefits of stimulus spending is weak. Studies suggesting that stimulus spending is effective assume that current economic activity does not affect fiscal policy. This assumption is questionable and, most important, the studies' results are very sensitive to it.

Alternatively, studies using a narrative approach based on historical and contemporary accounts of U.S. military buildups to identify government spending shocks find that the short-run effects of stimulus spending are small. These latter studies are grounded in historical reality and are thus more compelling.

Andrew T. Young is an associate professor of economics at West Virginia University and the College's BB&T Scholar. His papers have been published in Public Choice; Journal of Money, Credit and Banking; Review of Economics and Statistics; Managerial and Decision Economics; Review of Economic Dynamics; *and* Review of Austrian Economics.

CATO INSTITUTE

The return to fashion of fiscal stimulus was based more on wishful thinking than any new preponderance of evidence.

Introduction

In 1965 Milton Friedman was famously (or infamously, depending on your point of view) quoted as saying "We are all Keynesians now."[1] Though he later remarked that he had been taken out of context, Friedman accurately expressed a consensus among economists and policymakers at the time: countercyclical fiscal policy is an appropriate and effective means toward combating the business cycle. Specifically, in response to a recession, the federal government should increase its expenditures and/or provide tax relief to make up for the shortfall in the private sector's aggregate demand for goods and services. Nearly three decades after his publication of *The General Theory of Employment, Interest, and Money*, the policy prescriptions of Lord Keynes were dogma.

However, by the beginning of the new millennium all that had changed. Countercyclical fiscal policy had fallen into disrepute. A new consensus held that fiscal stimulus in the wake of a recession was politically infeasible and, even if it could be enacted in a timely fashion, stabilization policy was likely to be ineffective. Alan Blinder, the former vice chairman of the Federal Reserve's Board of Governors, went so far as to say that "virtually every contemporary discussion of stabilization policy by economists—whether it is abstract or concrete, theoretical or practical—is about monetary policy, not fiscal policy."[2]

But it would seem that the fashionableness of countercyclical fiscal policy is, well, countercyclical. Beginning in late 2007, recession and financial crisis led to a rehabilitation of Keynesian fiscal policy. Under the George W. Bush administration, $152 billion of tax rebates and other relief was passed in February of 2008. Almost exactly a year to the day later, President Barack Obama signed the $831 billion American Recovery and Reinvestment (ARRA) stimulus package into law.[3] Christina Romer, at that time the chair of Obama's Council of Economic Advisers, remarked that the ARRA was "simply the biggest, boldest countercyclical stimulus in American history" and that she "firmly believe[ed] that fiscal policy will have a crucial beneficial impact."[4] Pro-stimulus academic macroeconomists Antonio Fatás and Ilian Mihov declared that, "properly executed empirical studies support the view that a fiscal expansion will help the U.S. economy recover."[5] And even after nearly two years of continued unemployment above 9 percent, Nobel laureate Paul Krugman opined that the ARRA's only problem was that it "was much too small."[6]

But the return to fashion of fiscal stimulus was based more on wishful thinking than any new preponderance of evidence. National Bureau of Economic Research (NBER) macroeconomist Jonathan Parker noted that when the Obama "administration turned to economists—significantly academic economists—to help craft the size and details of the [ARRA] stimulus package" what they provided was "a wide range of answers . . . almost as useless as no answer." Indeed, the Cato Institute published a full-page *New York Times* ad arguing against government spending stimulus and carrying the signatures of over 200 economists, including three Nobel laureates.[7] Lacking any consensus from the experts, policymakers resorted to "using a multiequation macroeconomic forecasting model, one inconsistent with the best practice of modern macroeconomics."[8] Less than 15 years after Robert Lucas won the Nobel Prize in economics for forcefully arguing against the appropriateness of using such large scale macroeconometric models to design policy, the Obama administration placed an $800 billion bet that Lucas had been wrong.

This paper evaluates, particularly in regard to government spending, what confidence we should have in fiscal stimulus policies to restore the U.S. economy to robust expansion. My focus on spending is to make the subject manageable and also reflects the current administration's demonstrated preference for government spending relative to tax relief. As I will argue, there is *no consensus*

that fiscal stimulus spending is at all effective. Indeed, while one can find studies to support almost any position in the policy debate, the most compelling efforts to isolate exogenous changes in government spending and estimate their effect on gross domestic product (GDP) report that each additional dollar of stimulus yields less than one additional dollar in total GDP. In the terminology of economists: *the government spending multiplier is significantly less than one.* Increases in government expenditures are associated with decreases in private spending. The net change in total expenditures, then, falls short of the fiscal stimulus itself.

The debate over the appropriateness of Keynesian stimulus spending is ongoing, and the potential consequences are large. Despite virtually no prima facie evidence that the ARRA has been effective in restoring robust economic growth and full employment, the Obama administration is prepared to roll the dice again. Since September of 2011, the administration has been advocating an additional $447 billion of stimulus—about half of which would be new government expenditures—under the proposed "American Jobs Act."[9] With the U.S. federal government debt already standing at more than 70 percent of our GDP, we should demand critical evaluation of such a proposal from our policymakers.

In what follows I begin by putting recent fiscal stimulus spending in quantitative perspective relative to recent U.S. history. At least relative to recent U.S. experience, the ARRA was a very large spending package. Next, I discuss the intuition behind the concept of a government spending multiplier. Deciding on the appropriateness of fiscal stimulus hinges on estimating the size of this multiplier using U.S. data. I then discuss the problems involved in estimating the government spending multiplier. I review the existing estimates and conclude that while a wide range of estimates exist, the most plausible of those are significantly less than one. Last, I note that even advocates of fiscal stimulus, such as Fatás

and Mihov, admit that the long-run consequences for the national debt are often dire. Indeed, the findings of their own academic research suggest that the long-run costs are ominous. It is for these (negative) consequences that we truly have consensus amongst macroeconomists.

Recent Fiscal Stimulus in Perspective

Paul Krugman criticized the ARRA, and in particular its government spending component, for not being "big enough to do the job."[10] This certainly could be true and it cannot simply be dismissed. What I argue here is that, relative to recent U.S. experience, the new government spending initiated by the ARRA was quite large. This cannot rule out the possibility that countercyclical responses of U.S. government spending have been *in general* too small. However, it does move the discussion toward something more concrete than the otherwise vague characterizations of "too big" or "too small."

Figure 1 plots the quarterly percentage deviations of real (i.e., inflation-adjusted) U.S. federal government expenditures from their trend values. The basic data are 1947 to 2012, and a Hodrick–Prescott filter was used to compute a smoothly evolving trend.[11] To lend some historical perspective, vertical lines indicate the NBER beginning-of-recession dates and shaded areas correspond to the Korean War years as well as years of major U.S. involvement in the Vietnam War. A horizontal line at just over 40 percent corresponds to the peak deviation in the wake of the post-2008 fiscal stimulus policies. Outside of the large and positive deviations in federal spending during buildups for major military efforts, the increased spending that peaks in 2010 is extraordinary. Nothing like it has been previously seen during peacetime.

Additional perspective based on nonmilitary federal government expenditures is given in Figure 2. The ratio of federal nonmili-

The government spending multiplier is significantly less than one.

Figure 1
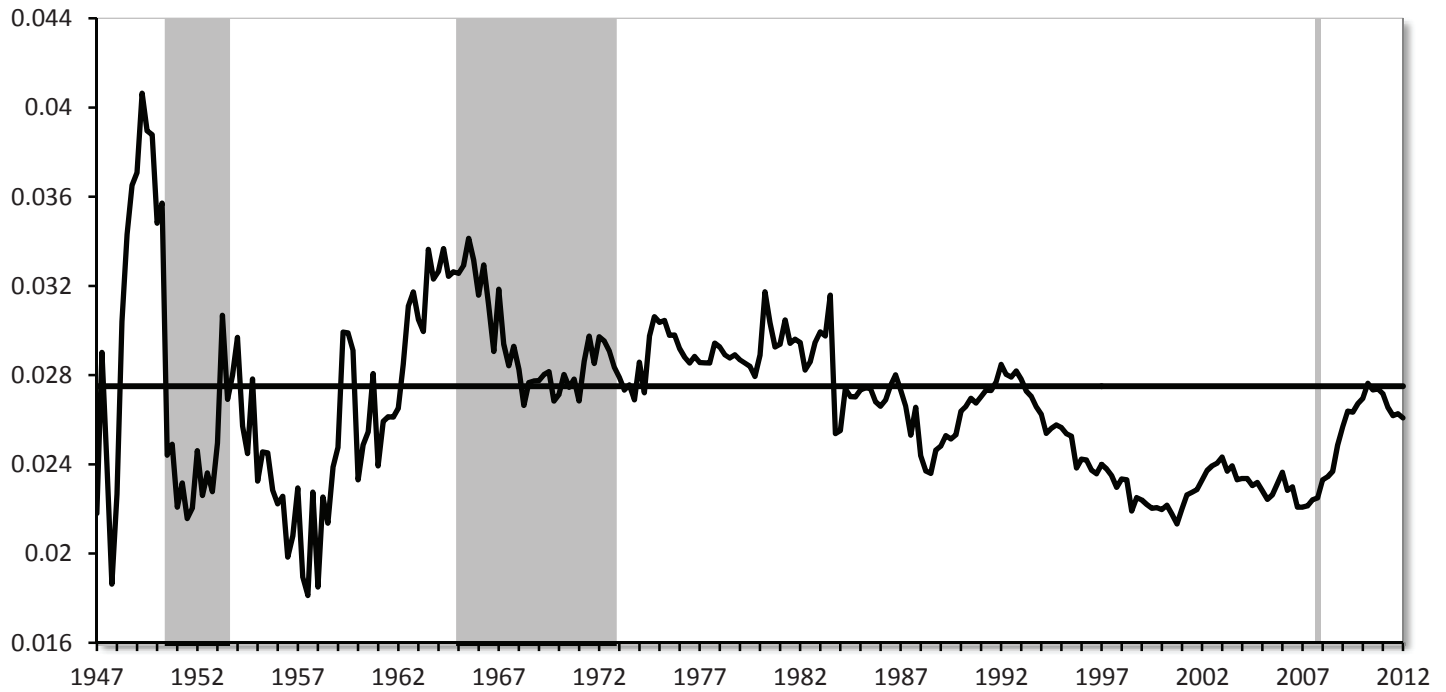**Deviations of Real Federal Government Expenditures from Trend, 1947–2012**
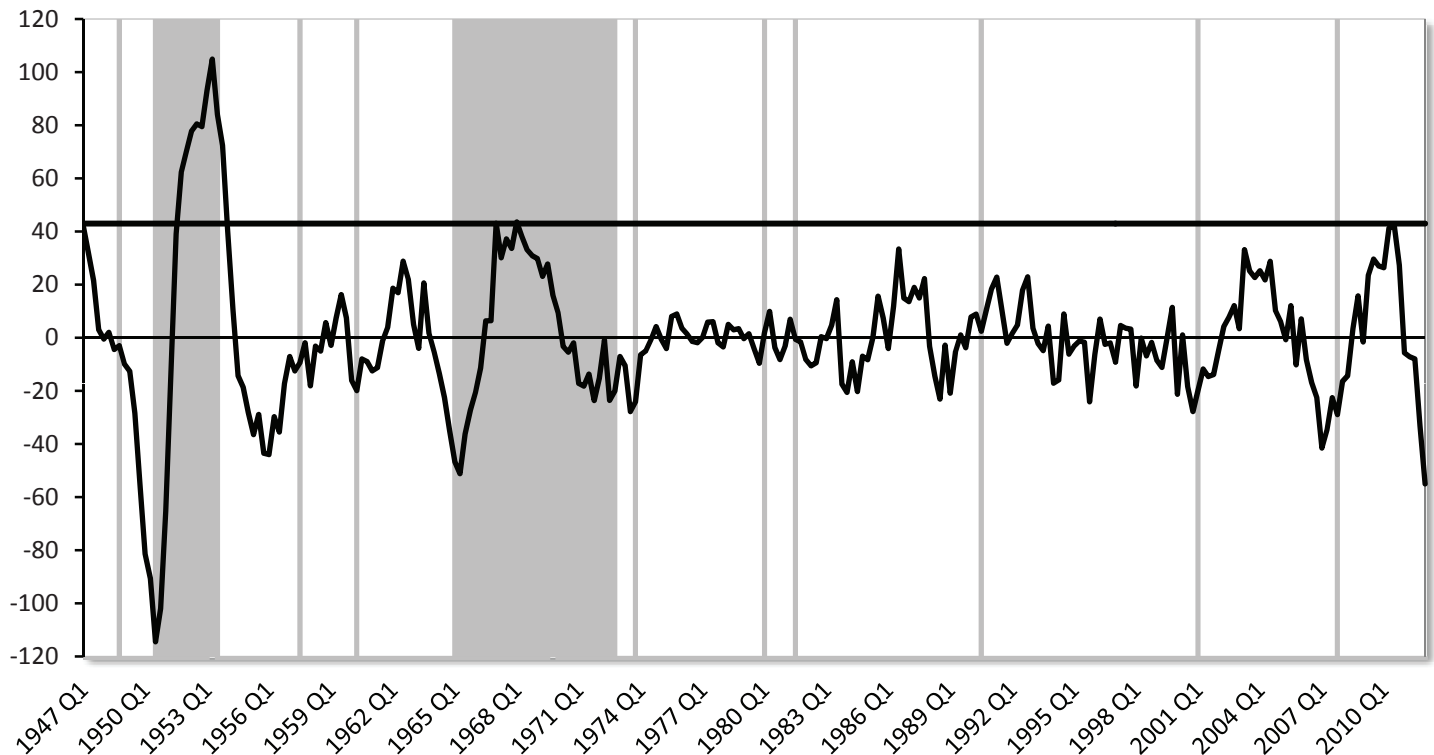
Source: U.S. Bureau of Economic Analysis. The federal component of government consumption expenditures, in billions of seasonally adjusted chained 2005 U.S. dollars. Observations are quarterly but stated at annual rates. Federal expenditures are detrended using the filter provided by Robert J. Hodrick and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking* 29, no. 2 (February 1997): 1–16. The smoothing parameter value is set according to Morten O. Ravn and Harald Uhlig, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics* 84, no. 5 (May 2002): 371–76. Shaded areas mark the Korean War and major U.S. involvement in the Vietnam War. Vertical lines mark National Bureau of Economic Research beginning-of-recession dates.

tary expenditures to GDP is plotted, again on a quarterly basis, from 1947 to 2012. The Korean War and Vietnam War eras are again shaded; a vertical line marks the beginning of the recent recession in the fourth quarter of 2007. The horizontal line in this case marks the 2010 peak of the series at over 2.7 percent of GDP. This is the highest since 1992. More important, measured from 2007, it is the largest single increase since the 1960s. In general, aside from the upward spike in 1992, nonmilitary expenditures as a percent of GDP had been trending downward since the 1970s. The ARRA and plans for additional federal stimulus, along with the loom-

ing entitlements associated with Obamacare may permanently reverse that trend.

From Figures 1 and 2 we can see that the fiscal stimulus in the wake of the Great Recession was remarkable. Federal government spending—even nonmilitary—has certainly been a larger share of the U.S. economy at times. However, in terms of fiscal stimulus policy what matters is the increase in expenditures. The federal government's fiscal response to the recession reversed basically all of the decreases in nondefense spending that had been achieved during the years of the Clinton administration and the "Contract with America" Congress.

Figure 2
Ratio of Federal Non-Military Expenditures to GDP, 1947–2012

Source: U.S. Bureau of Economic Analysis. The federal component of government consumption expenditures minus national defense expenditures divided by real gross domestic product, in billions of seasonally adjusted chained 2005 U.S. dollars. Observations are quarterly but stated at annual rates. Shaded areas mark the Korean War and major U.S. involvement in the Vietnam War. Vertical line marks National Bureau of Economic Research beginning-of-recession date in the fourth quarter of 2007.

Fiscal Multipliers

Of course, if you believe in the Keynesian model then a remarkable increase in government expenditures is precisely what the doctor ordered for the slumping U.S. economy. Furthermore, you hope that the effect on total spending in the economy is *even more remarkable* than the initial fiscal stimulus. The whole point of stimulus policy is to *stimulate* increases in total spending and incomes beyond the size of the stimulus package itself. In other words, the idea is for the initial injections into the economy to *multiply*. Whether the size of this effect—the *multiplier*—is large will depend on several factors, including individuals' marginal propensity to spend out of their incomes and their expectations of future tax liabilities, the extent to which deficit-financed government spending raises interest rates and crowds out private spending, and the amount of slack in the economy. These factors are not independent from one another. However, understanding each of their separate roles will facilitate thinking about why fiscal stimulus may or may not be effective.

Before considering each of the factors in turn, it will be helpful to place fiscal stimulus in the context of a simple framework. As would be common in an introductory macroeconomics course, think of an economy's total expenditures as the sum of three components: private consumption expenditures, private investment expenditures, and gov-

5

ernment expenditures. There are also expenditures associated with the import of foreign goods and the export of our own goods to foreigners, but we can have a meaningful discussion of fiscal policy without considering them explicitly. None of the fundamental ideas and intuitions below would be changed by doing so, and focusing on consumption, investment, and government expenditures keeps it simple.

The choices by individuals and policymakers that result in consumption, investment, and government expenditures are collectively referred to as the economy's aggregate demand for goods and services. The point of fiscal stimulus is to provide initial increases to aggregate demand. This is why such a policy to combat the business cycle if often referred to as *aggregate demand management*. The initial increases in aggregate demand are either directly from policymakers increasing the amount of government expenditures, or indirectly from their decreasing taxes. In the latter case, by decreasing the amount of taxes collected from individuals and businesses in the economy, net (or disposable) private incomes are higher. The hope is that individuals and businesses, finding themselves with more of their incomes and profits left over after paying taxes, increase their private consumption and investment expenditures.

Because the government has recourse to financial markets and can run deficits or (much less commonly in recent U.S. history) surpluses, a balanced-budget constraint is not binding when policymakers consider short-run fiscal policy changes. Policymakers can borrow to finance a deficit or use surplus revenues to buy financial assets (including the government's own outstanding debt).Therefore, increases or decreases in government spending and taxes can be contemplated and enacted independently of one another. For example, the ARRA provided for both increases in government spending and decreases in taxes. This, of course, would not have been possible if legislators were constrained by law or necessity to keep the federal government's budget balanced.

Increases or decreases in government spending and taxes can be contemplated and enacted independently of one another.

The Marginal Propensity to Expend

But the initial increases in expenditures need not be the whole story; indeed, they are almost certainly not. Any expenditure made in the economy generates income for some individuals. Those individuals then face a decision about what to do with the increase in their incomes. "To spend or not to spend?" That is the question. At the heart of multiplier effects, then, are individuals' marginal propensities to expend. The marginal propensity to expend is the percent of an additional (or marginal) dollar of income received that will be spent rather than saved.

Consider, for example, that the federal government undertakes a new highway maintenance program by spending $10 million. Government spending (and aggregate demand) immediately increases by $10 million. But when that money, it creates $10 million in new income for asphalt sellers, road crews, etc. These recipients of new income can save all or some of it, and then spend the rest. What they spend leads to a further increase in aggregate demand through an increase in private consumption expenditures and/or investment expenditures. If, for example, the marginal propensity to spend is 0.8, then individuals spend 80 percent of their new income—in this case, $8 million—and save the rest. In turn, that $8 million of new spending creates $8 million of new income for other individuals. If a marginal propensity to spend of 0.8 is typical, these other individuals also spend 80 percent of their new incomes, or $6.4 million. And the process continues as yet others receive the $6.4 million as income. With each round of new income generated, the dollar value of new expenditures falls because some part of that income—in this example, 20 percent—is saved rather than spent.

If one wants to know, for a given marginal propensity to expend, the total amount of new expenditures that can result from an initial fiscal stimulus, a bit of simple math is helpful. Denote the typical American's marginal propensity to spend by *MPS*. We can define the government spending multiplier

as $1/(1-MPS)$. It is called a multiplier because it tells us by how much some initial amount of new government expenditures (ΔG) can multiply in total new expenditures (ΔY). The process described in the previous paragraph can be expressed mathematically as

$$\Delta Y = (1/(1-MPS)) \times \Delta G.$$

Plug in the initial stimulus (ΔG) on the right hand side of the equation and, given people's tendency to spend out of additional income (MPS), one can easily solve for the total amount of new spending and incomes (ΔY) that can be generated.

Now, let's go back to our example where there is an initial stimulus in the form of $10 million of new government spending on highway maintenance. If the MPS is 0.8 then the government spending multiplier is five. Each dollar of fiscal stimulus can yield $5 of total increase in aggregate demand. A $10 million highway maintenance program will result in $50 million in total new expenditures and income. If, alternatively, the MPS is 0.9, then the multiplier is 10, and there is twice the bang for the stimulus buck. Intuitively, if at each round of the process, individuals spend a greater portion of their new incomes, then a greater portion of their new incomes subsequently becomes new income for additional individuals.

Note that for a given marginal propensity to expend, the government spending multiplier tells us how much a dollar of initial stimulus *can* yield in terms of total new expenditures; not necessarily how much it *will* yield. Most U.S. citizens spend more than 80 percent of the income that they receive. U.S. personal savings rates are actually quite low, so the marginal propensity to expend is likely to be quite high. However, not even the most ardent advocates of fiscal stimulus believe that the government spending multiplier is five or more. Typically, studies cited in support of fiscal stimulus report government spending multipliers above one but not more than two.[12] A multiplier of two would correspond to a marginal propen-

sity to spend of 0.5. Personal savings rates in the U.S. are nowhere near 50 percent, so for the empirical multiplier to be two or less there must be other factors associated with a stimulus that cause counteracting decreases in private consumption and investment expenditures.[13] In other words, there must be other important considerations not included in the simple Keynesian model.

Expectations of Future Tax Liabilities and Ricardian Equivalence

The federal government can certainly increase its expenditures beyond its revenues in the short run, but this does not mean that the bills will not come due. The government runs spending deficits by borrowing; it sells its treasury securities on financial markets to U.S. and foreign individuals. Those treasuries represent promises of future payments from the federal government (amounting to an outstanding total of about $16 trillion as of late). The government will eventually have to raise revenue with a present value equal to those liabilities.

Macroeconomist Robert Barro put forth the argument in the 1970s that government spending multipliers will be close to zero precisely because when government runs a deficit, taxpayers will rationally add the value of that deficit to their expected future tax liabilities.[14] In anticipation of having to make good on these liabilities, they increase their savings, which implies that they *decrease* their expenditures. If the decrease in private expenditures exactly matches the new government expenditures, then the negative and positive multiplier effects will exactly offset. The multiplier will be zero. For every deficit-financed dollar of government expenditures, taxpayers hold a dollar back in anticipation of having to pay higher taxes in the future to retire the new government debt. In a reference to the classical economist David Ricardo, who in the 1800s discussed similar insights in relation to British war bonds, the hypothesis that taxpayers will internalize government's intertemporal budget constraint is called Ricardian equivalence.

If the decrease in private expenditures exactly matches the new government expenditures, then the negative and positive multiplier effects will exactly offset. The multiplier will be zero.

Just as even the most ardent advocates of Keynesian stimulus do not believe that the spending multiplier is five, not even Robert Barro claims that the short-run government spending multiplier is zero. A year after the ARRA, Barro stated that his best reading of the evidence and recent experience suggests a multiplier of 0.4 in the first year and 0.6 over two years.[15] A multiplier significantly less than one means that every $1 of government stimulus buys the economy less than a $1 increase in total income. If the spending multiplier is 0.6, for example, the ARRA's approximately $500 billion in new spending bought only about $300 billion in total new expenditures and income; if the multiplier is 0.4, only about $200 billion. One doesn't need to be an economist to see how that's a raw deal.

Crowding Out

To believe that there may be spending deflators, rather than multipliers, associated with stimulus spending, one does not have to believe that U.S. citizens are experts on public finance, poring over Congressional Budget Office projections and forming rational expectations of their future tax liabilities. Largely one just has to believe that they respond to market prices and, in particular, market interest rates.

When the federal government pursues spending in excess of its current tax revenues, it has to turn to financial markets and compete with private borrowers for funds. The increased demand for funds will, all else equal, put upward pressure on interest rates, making borrowing more costly. By raising the cost to private borrowers—both individuals and businesses—government deficit spending tends to crowd out private investment expenditures and consumption expenditures that are sensitive to interest rates (e.g., car purchases).

Furthermore, higher costs of borrowing also mean higher returns to saving. As the government increases its demand for funds, this must be satisfied by increases in private savings. This also discourages private consumption expenditures.

Targeting Stimulus to Where There's "Slack"

Another factor upon which the effectiveness of fiscal stimulus depends is the amount of "slack" in the economy where it is targeted. This is acknowledged in the standard textbook exposition of the Keynesian model where fiscal policy is conceived of as shifting aggregate demand along the short-run aggregate supply curve (SRAS) (Figure 3). The SRAS curve represents firms' willingness to produce various amounts of goods and services (Y) at various levels of prices (P). As the economy approaches full employment and its potential level of output, the slope of the SRAS curve increases, approaching vertical. This implies that when the economy is running on all cylinders increases in aggregate demand will largely translate into rising prices (inflation) rather than higher production. Alternatively, the flatter portion of SRAS represents levels of (Y) considerably below potential. There is considerable slack in terms of the available capital and labor, and increases in aggregate demand can prime the pump and bring those resources into use in production of good and services.

Obviously the ARRA was enacted during a time of considerable slack in the economy. Unemployment peaked at 10 percent in July of 2009 and remained (and remains) stubbornly above 8 percent. However, the recession did not hit all areas of the country equally. Considering the allocation of funds across U.S. states and based on available measures of slack (e.g., unemployment rates and state GDP levels), my own research with Russ Sobel indicates that the ARRA has not been consistently targeted to areas operating below potential. (Nor, for that matter, was it targeted to states where the marginal propensity to expend is particularly high.) Rather, political variables, such as U.S. House of Representatives committee memberships, were better predictors of funding levels from the stimulus. And the strongest and most robust predictor of a state's ARRA allocation is the ratio of a state's federal grants and payments to its federal taxes paid previous

Political variables, such as U.S. House of Representatives committee memberships, were better predictors of funding levels from the stimulus.

Figure 3
The Effect of Stimulus in the Keynesian SRAS/AD Mode Is Contingent on How Close the Economy Operates to Its Potential
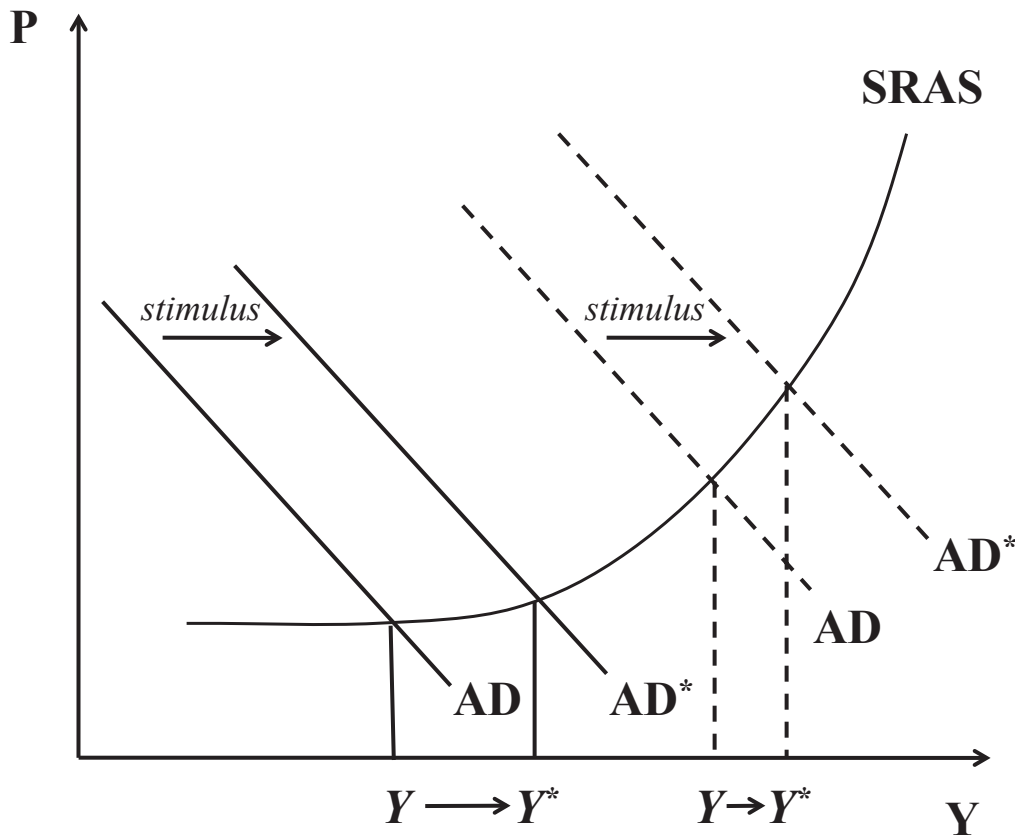
Source: Author's creation; standard aggregate supply/aggregate demand graph.

to the stimulus. In other words, states that previously captured large amounts of federal funds also managed to capture larger portions of the ARRA stimulus.[16]

While my research with Sobel does not speak directly to the effectiveness of fiscal stimulus, it serves as a reminder that fiscal policy is never a matter of simply increasing spending or decreasing taxes. In the real economy, stimulus packages are products of the political process and objects of the standard rent-seeking games that politicians and special interests play. This not only brings into question whether stimulus spending will be well-targeted to areas of slack in the economy. It should also lead us to question

the non-stimulus value of the investments that are being made and services that are being provided. Even if a rush to find "shovel-ready" projects does not itself lead to unwise expenditures, the political competition to capture funds and please constituents is likely to.

The (Lack of) Evidence

Finally we come to the big question: Does Keynesian fiscal stimulus actually work? Is the spending multiplier greater than one? My own reading of the literature draws five general conclusions:

In the real economy, stimulus packages are products of the political process.

Even though the U.S. economy has not roared back following the American Recovery and Reinvestment Act, one might argue that it would be doing far worse now had it not been for the federal stimulus spending.

- There are numerous studies finding that the U.S. government spending multiplier is greater than one; there are also numerous studies finding that it is less than one.
- The primary challenge tackled by all of these studies is to identify exogenous changes in government spending and then trace out their effects on total spending over time.
- The results are very sensitive to how these exogenous government spending changes (or shocks) are identified.
- The most plausible identification strategies—based on increases in government spending associated with military buildups—result in multiplier estimates that are less than one.
- Even researchers who report large spending multipliers acknowledge that the long-run costs of deficit-financed fiscal stimulus are likely to be large.

At best, our confidence in the effectiveness of fiscal stimulus should be small relative to the price tags attached to a package like the ARRA. At worst, I will argue contra Antonio Fatás and Ilian Mihov that studies with the greatest claim to being properly executed suggest that a stimulus' bang is quite a bit smaller than its buck.

Identifying Exogenous Changes in Spending

Even though the U.S. economy has not roared back following the ARRA, one might argue that it would be doing far worse now had it not been for the federal stimulus spending. Fiscal stimulus may get us a period of weak growth and employment in exchange for one where we plunge into a deep depression. Alternatively, if stimulus packages tend to be passed late in a recession—or after the recession has actually ended—then they will tend to be positively correlated with growth and rising employment. In either case, we will not be getting a clear picture of the macroeconomic effects of the stimulus itself.

Both of the above scenarios arise because fiscal stimulus is what economists refer to as an endogenous event. Stimulus funds do not randomly fall onto the economy like so much manna from heaven. Rather, they are the result of legislation that is designed because of and in response to macroeconomic conditions. We cannot just look at subsequent economic events and determine that they were caused by a stimulus, because a stimulus is itself caused by the unfolding chain of economic events.

What researchers would like to identify are exogenous changes in government spending—changes that were not caused by unfolding events in the economy. Economists refer to such changes as shocks since they arise from outside of the specific system of variables that one wants to analyze. Tracing out the effects of such shocks to government spending is the ideal way to estimate the government spending multiplier. But identifying such shocks is easier said than done. Fiscal stimulus is countercyclical policy; it is pretty much by definition endogenous to the business cycle, and so are variations in spending associated with automatic stabilizers such as unemployment insurance that, by design, increase their payments when economic activity falls. Other legislated changes in government spending are also generally influenced by economic conditions.

Economists who study the effects of fiscal stimulus have addressed this challenging problem using one of two approaches (or a combination of the two). The first approach involves the estimation of structural vector autoregressions (SVARs). A vector autoregression (VAR) is a system of equations that describes how certain economic variables (e.g., GDP and government spending) evolve over time and affect one another. In regard to the government spending multiplier, a researcher will be interested in using macroeconomic data to estimate a VAR and quantify the effects of a shock to government spending on GDP over time.

The challenge in working with VARs is to identify the shocks (or exogenous changes)

to government spending. A researcher cannot simply examine all changes to government spending. Many of those spending changes will be the result of (at least in part) changes in economic activity. To the extent that the government spending changes are associated with changes in GDP over time, the researcher will be unable to tell to what extent those GDP changes are caused by the government spending rather than merely associated with previous changes in economic activity. The researcher wants to identify exogenous shocks to government spending so that he or she can be confident that the estimated effects on GDP are the result of the stimulus.

To solve this identification problem, a researcher must impose additional structure on the VAR (hence, an SVAR). A common strategy is to order the shocks of the different variables included in the SVAR. Specifically, a researcher will assume that shocks to government spending can affect economic activity and GDP immediately, but not vice versa. The technical details about the estimation of SVARs would be inappropriate here, but a simple example is provided in the appendix of this paper for the interested reader. What is important to note is that the results of estimating a SVAR are very sensitive to the particular ordering that one assumes.

As means of illustration, I estimated a simple bivariate SVAR in U.S. real GDP and federal government spending. I included both variables in equations containing six lagged values of each. The data I used to estimate the system is quarterly data for both GDP and government expenditures from 1947 Q3 through 2012 Q1, and I took the natural logs of both series, as is standard in most studies. Figure 4 presents the resulting *impulse responses* of U.S. real GDP to a federal government spending shock. An impulse response plots how a shock to one variable affects another variable in the system over time. The impulse response in panel (a) is computed assuming that government spending is ordered first; the one in panel (b) is computed

assuming, instead, that real GDP is ordered first. The difference is like night versus day. In panel (a), a shock to government spending has positive effects on GDP for a full year. In panel (b), the effect of a federal spending shock on GDP is immediately zero and then actually turns negative after another quarter passes.

Numerous papers estimating the government spending multiplier using U.S. data employ a SVAR approach based on the ordering and placement of other restrictions on the shocks. For example, papers by Antonio Fatás and Ilian Mihov in 2001and Olivier Blanchard and Roberto Perotti in 2002 assume that government spending is "predetermined with respect to macroeconomic shocks." In other words, the shock to government spending is ordered first; in a given quarter only it, and not exogenous changes to GDP, will increase or decrease government spending. This is precisely the type of assumption that I made in computing the panel (a) impulse response where GDP responds positively to a government spending shock. According to Fatás and Mihov, this assumption is a "plausible, but unfortunately untestable hypothesis."[17] They report that the U.S. government spending multiplier is greater than one. A 2007 paper by macroeconomists Jordí Gali, Javier Vallés, and J. David López-Salido echoed these earlier findings, reporting that the government spending multiplier reaches 1.74 two years after a spending shock.[18]

My own little SVAR analysis (Figure 4) is a naive one and, admittedly, does little justice to the published studies by other scholars. For instance, other variables in addition to government spending and GDP should be included in any serious study of the effects of fiscal stimulus. However, my simple example highlights the point I wish to make here: the results are very sensitive to the structure that one chooses to impose on the SVAR. And while the predetermined nature of government spending certainly is untestable, there are reasons to doubt whether or not it is really all that plausible. First and

The results of estimating a structural vector autoregression are very sensitive to the particular ordering that one assumes.

Figure 4
Estimated Effects of Real GDP to a Federal Government Expenditures Shock

(a)

(b)
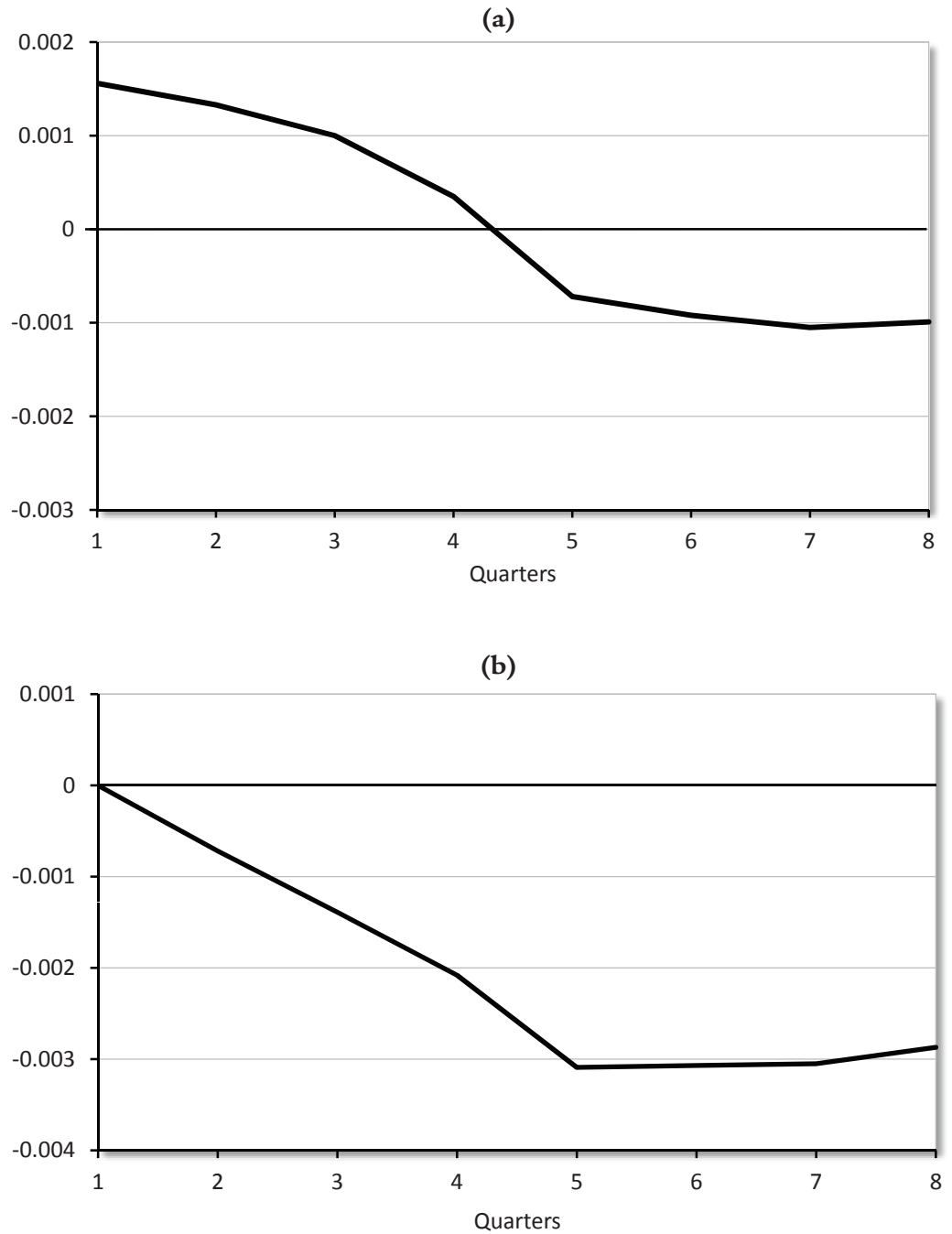
Source: Source: Author's calculations.
Note: Both impulse responses are based on a VAR incorporating quarterly real GDP and federal government expenditures from the BEA for 1948 Q3 to 2012 Q1. Six are lags are assumed for each equation and natural logs are taken of each variable. In the left panel (a) federal spending is ordered prior to real GDP; in the right panel (b) real GDP is ordered first. Depending on which restriction is assumed, the estimated response of real GDP to a federal spending shock is (a) positive for an entire year or (b) always negative. The vertical axis measures the change in the natural log of real GDP to a standard deviation innovation to the natural log of federal government expenditures.

foremost, it assumes that policymakers have a lack of access to, or an ignorance of, economic data that are being collected throughout a given quarter. (This seems especially implausible when using data samples that overlap the Internet age.) Also, in any SVAR study there are numerous variables associated with economic activity that are left out of the analysis. If those variables affect both GDP and policymakers' behavior in a given time period, then shocks that are identified as directly occurring to GDP may also affect government spending. For example, assume that economic conditions deteriorate in the European Union. This could certainly impinge on U.S. economic activity, and U.S. policymakers are cognizant of this; their legislative activities could then be directly influenced by the overseas developments.

When researchers abandon the predetermined government spending assumption in favor of alternative identification strategies, the results are drastically different. For example, in a 2009 paper Andrew Mountford and Harald Uhlig identified government spending shocks as being associated with increases in spending for at least a year and business cycle shocks as ones that caused government revenue, GDP, consumption, and investment to move in the same direction. They report a U.S. government spending multiplier of merely 0.07 after two years.[19] (The largest effect that they report is in the first year, and even then, the spending multiplier is only 0.65.)

Evidence Based on Variation in Military Spending

An alternative approach to identifying government spending shocks was pioneered by Valerie Ramey and Matthew Shapiro in a 1998 paper. They adopted a narrative approach and use historical accounts and contemporary *Business Week* stories to isolate episodes of post–World War II U.S. military buildups. As Ramey and Shapiro explain, "military buildups occur rapidly and unexpectedly, so they are naturally modeled as shocks[;] they are driven by imperatives of foreign policy. . . . [B]ecause they are driven by geopolitical shocks, military buildups are likely to be exogenous with respect to macroeconomic variables." [20]

Researchers following the Ramey and Shapiro approach report modest spending multipliers. Furthermore, they find that private consumption expenditures tend to fall in the wake of a government spending shock.[21] In a 2008 paper, Valerie Ramey reports that the post–World War II U.S. spending multiplier is between 0.6 and 0.8.[22] Using the most generous estimate, this implies that for each $1 of federal stimulus spending, $0.20 of private expenditures is crowded out, and the total increase in GDP will be less than the new government expenditures. The least generous estimate implies that the total increase in GDP is only 60 percent of the new government expenditures.

The narrative approach based on military buildups is by no means a perfect identification strategy, but it is in at least two ways appealing relative to the SVAR approach. First, shocks to government spending become grounded in historical reality. They are not simply fluctuations in spending that are identified when an SVAR is estimated; they rather correspond to historical episodes where the motives of policymakers can be described, albeit imperfectly. Second, advocates of federal stimulus spending often express a "go big or go home" approach. Military buildups are large increases in government spending that are comparable to fiscal stimulus packages.

In the Long Run We Are All in Debt

The evidence on the size of the U.S. government spending multiplier is mixed, at best, and I have argued above that the most plausible evidence suggests a multiplier significantly less than one. Meanwhile, the severity of the long-term consequences of the ARRA seems clear. Figure 5 demonstrates the massive increase in federal government debt held by the public as a percent of U.S. GDP. In 2007 Q4, this ratio stood at about 0.36. By 2012 the federal debt was more than 70 per-

Using the most generous estimate, this implies that for each $1 of federal stimulus spending, $0.20 of private expenditures is crowded out, and the total increase in GDP will be less than the new government expenditures.

Figure 5
Ratio of Federal Debt Held by the Public to GDP, 1970–2012
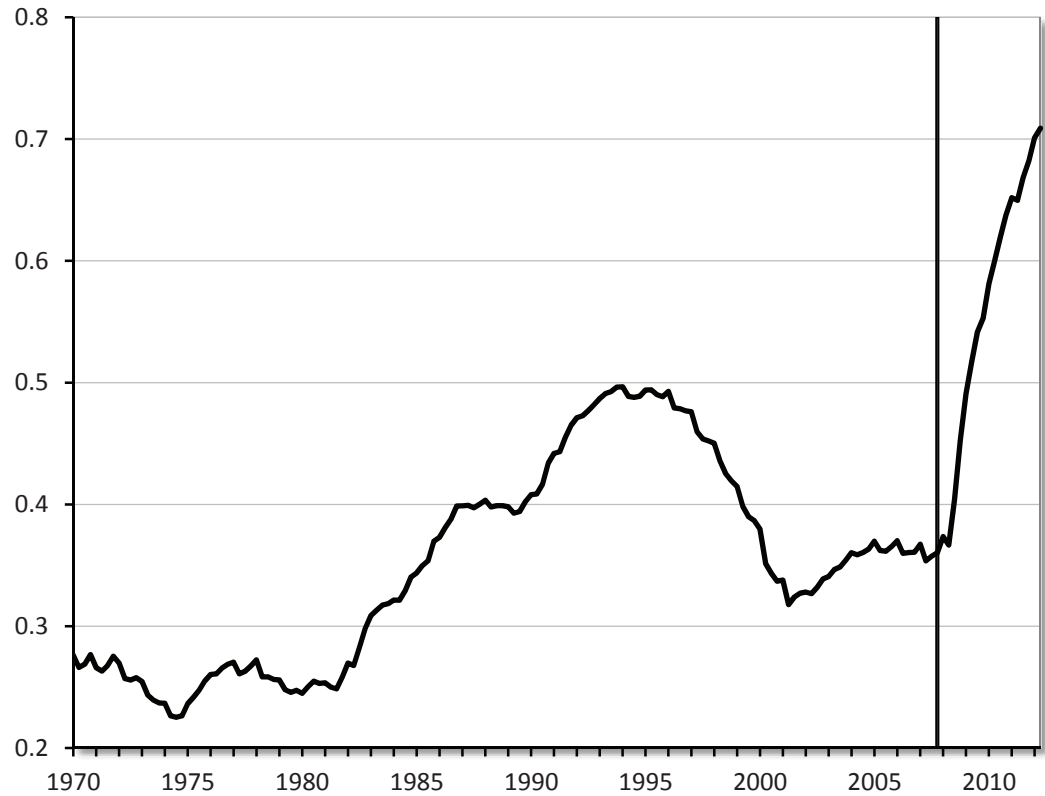
Source: Federal Reserve Bank of St. Louis Federal Reserve Economic Database, using federal debt held by the public divided by nominal GDP. Vertical line corresponds to 2007 Q4, the beginning of the Great Recession according to National Bureau of Economic Research business cycle dates.

> **Pro-stimulus economists Antonio Fatás and Ilian Mihov report that "governments that use fiscal policy aggressively induce significant macroeconomic instability."**

cent of the U.S. GDP. The debt-to-GDP ratio has almost doubled in less than five years; it has increased by more than 160 percent since 1970. The ARRA is clearly not the only cause of the debt explosion. For example, another important contributor is increasing Medicare expenditures. However, the $800 billion plus ARRA has exacerbated the problem, as will any of the proposed subsequent stimulus spending.

Whatever its sources, all of this debt entails future tax liabilities. While some may doubt that U.S. citizens always internalize the federal government's long-run budget constraint, it is also hard to believe that such large and looming liabilities will not serve to dampen private economic activity. Furthermore, there is considerable uncertainty as to

exactly how and when the necessary future revenues will be raised. Uncertainty is anathema to vibrant, future-oriented business activity.

Remarkably, in a 2003 study of data from 91 countries, pro-stimulus economists Antonio Fatás and Ilian Mihov report that "governments that use fiscal policy aggressively induce significant macroeconomic instability."[23] Even more remarkably, looking back on that research in a 2009 paper titled "Why Fiscal Stimulus is Likely to Work," they state that "changes in fiscal policy, present and future, will create instability in the macroeconomic environment that can lead back to a lack of confidence, uncertainty and hurt investment and growth. . . . But our view is that the benefits of a well-executed short-term

14

fiscal stimulus are larger than the long-term costs."[24] Their sanguinity concerning the size of the short-run government spending multiplier is unjustified given the existing evidence, especially given that no one seriously disputes the long-run burden that legislation like the ARRA places on the economy.

Conclusion

Why in the world are we all Keynesians again? The stimulus spending of the American Recovery Reinvestment Act has been an apparent failure. The economy has not returned to a path of robust economic growth. Unemployment has remained stubbornly above 8 percent. Yet many pundits, policymakers, and academics continue to encourage additional increases in government spending. For some, like Nobel laureate Paul Krugman, the only problem with the ARRA is that even more than $800 billion was simply not large enough. Their solution, therefore, is to argue for additional and larger stimulus packages.

But the evidence on the effectiveness of stimulus spending is weak. Studies finding a government spending multiplier greater than one typically rely on identifying government spending shocks by assuming that government spending does not respond to contemporaneous changes in economic activity. However, if that assumption is incorrect, then those researchers are attributing independent changes in GDP to changes in government spending. Data on underlying activity is readily available to policymakers within a given quarter. Also, factors such as changes in other countries' economic conditions can easily influence both domestic economic activity and fiscal policy decisions. For both reasons, the critical assumption of these studies is likely to be violated.

I have argued above that studies identifying shocks with the changes in government spending associated with U.S. military build-ups are more plausible. These studies report government spending multipliers less than

one. They find that private consumption expenditures fall in the wake of a shock. This is consistent with the insights of Robert Barro concerning Ricardian equivalence. Individuals in the economy internalize the government's intertemporal budget constraint. Responding to new future tax liabilities and higher interest rates, individuals offset new government spending, at least in part, by decreasing their own private expenditures.

In the above discussion, I have focused entirely on government spending as opposed to tax cuts. This is largely because both the ARRA and the policy proposals of the Obama administration moving forward have been based on the premise that the federal government can spend its way out of the economic malaise. Also, the same structural vector autoregression studies that report a government spending multiplier greater than one find that the U.S. tax multiplier is less than one.[25]

However, a 2010 paper by Christina Romer and David Romer reports a tax multiplier of nearly three.[26] Romer and Romer pursue a narrative approach to identifying tax shocks that is similar in spirit to Ramey and Shapiro's for identifying spending shocks. Romer and Romer utilize U.S. presidential speeches and congressional reports to provide evidence of tax changes that were not motivated by changes in macroeconomic conditions. (Given that Christina Romer is on record as an enthusiastic architect and proponent of the ARRA, it is interesting to note that the paper's discussion of its findings refers only to tax increases and their contractionary effects; never the possible stimulating effects of tax cuts.) To the extent that the narrative approach is more compelling than SVARs, Romer and Romer's evidence suggests that tax cuts are relatively more effective as fiscal stimulus than are government spending increases.

Stimulus packages like the ARRA come with staggering price tags, and even their proponents admit that they result in public debt burdens that are likely to dampen growth in the future and destabilize the

The evidence on the effectiveness of stimulus spending is weak.

15

It is hard to justify stimulus spending packages in light of the almost certain and large long-run costs.

macroeconomy. Given weak evidence of significant short-run benefits, it is hard to justify stimulus spending packages in light of the almost certain and large long-run costs. In September of 2011, the Obama administration advocated an additional $447 billion of stimulus under the proposed "American Jobs Act." The large majority of this proposed legislation was blocked by House Republicans.[27] However, the continued stagnation of U.S. economic growth and employment, along with the reelection of President Obama, suggests that further stimulus spending could be on the horizon. With federal government debt already in excess of 70 percent of GDP, we should insist that our policymakers evaluate such proposals critically and, lacking solid evidence for their effectiveness, put the federal checkbook away.

Appendix:
Details of a Simple VAR

A simple two-variable VAR might take the form,

$$GDP_t = \alpha_0 + \alpha_1 GDP_{t-1} + \alpha_2 GDP_{t-1} + \alpha_3 GOVT_{t-1} + \alpha_4 GOVT_{t-2} + \varepsilon_{1,t}$$
$$GOVT_t = \beta_0 + \beta_1 GDP_{t-1} + \beta_2 GDP_{t-1} + \beta_3 GOVT_{t-1} + \beta_4 GOVT_{t-2} + \varepsilon_{2,t}$$

,

where the t indexes time periods (say quarters) so that over time, policy $GOVT$ affects economic activity GDP and the level of economic activity also affects policy.

A researcher wants to estimate the parameter values (i.e., the αs and the βs) to understand the patterns and magnitudes of those interactions. Then, to understand how a shock to $GOVT$ affects the system, one has to first identify those shocks, and that involves imposing restrictions on the covariance structure between the components of the system's error terms, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$. The researcher presumes that the components of $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are the shocks to both GDP and $GOVT$.

The simplest way to identify these underlying shocks is to assume a specific ordering of those shocks. Call the underlying shocks to GDP and $GOVT$, respectively, $v_{gdp,t}$ and $v_{govt,t}$. One ordering would assume that shocks to government spending can affect economic activity immediately, but not vice versa:

$$\varepsilon_{1,t} = v_{gdp,t} + \theta v_{govt,t} \quad \text{and} \quad \varepsilon_{2,t} = v_{govt,t}.$$

As assumed above, the shock to government spending $v_{govt,t}$ has an impact on both the GDP and $GOVT$ equation directly. The parameter θ determines how strongly a shock to $GOVT$ initially impacts GDP. On the other hand, the shock to economic activity $v_{gdp,t}$ only enters the GDP equation directly and it affects $GOVT$ only indirectly through its effect on GDP over time. Alternatively, one could order GDP first by assuming that $\varepsilon_{1,t} = v_{gdp,t}$ and $\varepsilon_{2,t} = v_{govt,t} + \theta v_{gdp,t}$.

In Figure 4 in the text, the panel (a) impulse response is computed based on the assumption that $\varepsilon_{1,t} = v_{gdp,t} + \theta v_{govt,t}$ and $\varepsilon_{2,t} = v_{govt,t}$. The panel (b) impulse response is computed based on the assumption that $\varepsilon_{1,t} = v_{gdp,t}$ and $\varepsilon_{2,t} = v_{govt,t} + \theta v_{gdp,t}$.

Notes

1. "The Economy: We are All Keynesians Now," *Time*, December 31, 1965.

2. Alan S. Blinder, "The Case Against Discretionary Fiscal Policy," CEPS Working Paper no. 100 (June 2004).

3. The original estimate of the total tax cuts and expenditure increases of the ARRA was $787 billion. The Congressional Budget Office has since increased that estimate: http://www.cbo.gov/sites/default/files/cbofiles/attachments/02-22-ARRA.pdf. In a September 16, 2012, *Washington Post* piece, Robert Samuelson also argued that "the true stimulus also includes subsequent tax cuts and spending increases plus 'automatic stabilizers.' These refer to the budget's tendency to swing into deficit during a recession because tax revenues fall and spending on unemployment benefits and other safety-net programs rise. Budget deficits broadly measure stimulus. From 2009 to 2012, they're about $5.1 trillion." See Robert Samuelson, "Bernanke on the Brink," *Washington Post*, http://www.washingtonpost.com/opinions/robert-j-samuelson-bernanke-on-the-brink/2012/09/16/9c94a66c-fe93-11e1-8adc-499661afe377_story.html.

4. See Christina Romer's comments at the National Association of Business Economists' 25th annual Washington Economic Policy Conference, March 3, 2009, at http://www.whitehouse.gov/administration/eop/cea/speeches-testimony/03032009/.

5. Antonio Fatás and Ilian Mihov, "Why Fiscal Stimulus is Likely to Work," *International Finance* 12, no. 1 (Spring 2009): 70.

6. Paul Krugman, "Keynes Was Right," *New York Times*, http://www.nytimes.com/2011/12/30/opinion/keynes-was-right.html.

7. David Boaz, "Economists against the Stimulus," *Cato at Liberty* (blog), http://www.cato-at-liberty.org/economists-against-the-stimulus/.

8. Jonathan A. Parker, "On Measuring the Effects of Fiscal Policy in Recessions," *Journal of Economic Literature* 49, no. 3 (September 2011): 704.

9. Janet Novack, "Obama Proposes $447 Billion 'Jobs' Package," *Forbes*, http://www.forbes.com/sites/janetnovack/2011/09/08/obama-proposes-447-billion-jobs-package/.

10. Paul Krugman, "The Conscience of a Liberal," *New York Times*, http://krugman.blogs.nytimes.com/2011/09/05/on-the-inadequacy-of-the-stimulus/.

11. The Hodrick-Prescott filter is a popular method for fitting a smooth trend through macroeconomic data. See Robert J. Hodrick and Edward C. Prescott, "Postwar U.S. Business Cycles: Am Empirical Investigation," *Journal of Money, Credit, and Banking* 29, no. 1 (February 1997): 1–16.

12. For example, see Antonio Fatás and Ilian Mihov, "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," CEPR Discussion Paper no. 2760 (April 2001); and Olivier Blanchard and Roberto Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics* 117, no. 4 (November 2002): 1329–68.

13. Strictly speaking, looking at data on personal savings rates indicates (the inverse of) Americans' *average* propensity to expend; not their *marginal* propensity to expend. But the difference between observed savings rates and the marginal propensity to expend necessary to imply a multiplier of two (or less) is too large to not suggest other factors that must be taken into account.

14. Robert J. Barro, "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82, no. 6 (November–December 1974): 1095–117.

15. Robert J. Barro, "The Stimulus Evidence One Year On," *Wall Street Journal*, February 23, 2010, http://scholar.harvard.edu/barro/files/wsj_10_0223_stimuluspackage.pdf.

16. Andrew T. Young and Russell S. Sobel, "Recovery and Reinvestment Act Spending at the State Level: Keynesian Stimulus or Distributive Politics," *Public Choice* (forthcoming).

17. See Olivier Blanchard and Roberto Perotti, "An Empirical Characterization of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics* 117, no. 4 (November 2002): 1329–68; and Antonio Fatás and Ilian Mihov, "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," CEPR Discussion Papers no. 2760 (April 2001). The quotes are from Fatás and Mihov, p. 6.

18. Jordí Gali, Javier Vallés, and J. David López-Salido, "Understanding the Effects of Government Consumption Spending on Consumption," *Journal of the European Economic Association* 5, no. 4 (March 2007): 227–70.

19. See Andrew Mountford and Harald Uhlig, "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 24, no. 6 (April 2009): 960–92.

20. Valerie A. Ramey and Matthew D. Shapiro,

"Costly Capital Reallocation and the Effects of Government Spending," *Carnegie-Rochester Conference Series on Public Policy* 48, no. 1 (June 1998): 145–94. Quote is from p. 174.

21.   Recent papers include Wendy Edelberg, Martin Eichenbaum, and Jonas D. M. Fisher, "Understanding the Effects of a Shock to Government Purchases," *Review of Economic Dynamics* 2, no. 1 (January 1999): 166–206; and Craig Burnside, Martin Eichenbaum, and Jonas D.M. Fisher, "Fiscal Shocks and Their Consequences," *Journal of Economic Theory* 115, no. 1 (March 2004): 89–117.

22. Valerie Ramey, "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics* 126, no. 1 (February 2011).

23.   Antonio Fatás and Ilian Mihov, "The Case for Restricting Fiscal Policy Discretion," *Quarterly Journal of Economics* 118, no. 4 (November 2003): 1419–47.

24.   Antonio Fatás and Ilian Mihov, "Why Fiscal Stimulus is Likely to Work," *International Finance* 12, no. 1 (May 2009): 57–73.

25.   For example, Blanchard and Perotti (2002) report a tax multiplier around 0.8. (See note 14.)

26.   Christina Romer and David Romer, "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review* 100, no. 3: 763–81. Mountford and Uhlig, using their alternative SVAR identification strategy, also report a tax multiplier of 1.63 after two years, which is considerably larger than the spending multiplier of 0.07 that they report. (See note 21.)

27.   An exception being the Jump Start Our Business Start-Ups (JOBS) Act, signed into law in April of 2012. JOBS is a package of incentives for small businesses; it is largely devoid of ARRA-type fiscal stimulus.

RELATED PUBLICATIONS FROM THE CATO INSTITUTE

Should U.S. Fiscal Policy Address Slow Growth or the Debt? A Nondilemma by Jeffrey Miron, Cato Institute Policy Analysis no. 718 (January 8, 2013)

Fiscal and Generational Imbalances and Generational Accounts: A 2012 Update by Jagadeesh Gokhale, Cato Institute Working Paper no. 12 (November 13, 2012)

Fiscal Policy Report Card on America's Governors: 2012 by Chris Edwards, Cato Institute White Paper (October 9, 2012)

The Misuse of Top 1 Percent Income Shares as a Measure of Inequality by Alan Reynolds, Cato Institute Working Paper no. 9 (October 4, 2012)

World Hyperinflations by Steve H. Hanke and Nicholas Krus, Cato Institute Working Paper no. 8 (August 15, 2012)

What Made the Financial Crisis Systemic? by Patric H. Hendershott and Kevin Villani, Cato Institute Policy Analysis no. 693 (March 6, 2012)

RECENT STUDIES IN THE
CATO INSTITUTE POLICY ANALYSIS SERIES