Why Do Economists Still Disagree over Government Spending Multipliers?

Daniel Carroll

Public debate about the effects of government spending heated up after record-large stimulus packages were enacted to address the fallout of the financial crisis. Almost as noticeable as the discord was the absence of consensus among prominent economists on the issue. While it seems a simple problem to estimate the effect of government spending on output—the size of the government multiplier—it is anything but.

Over the past several years, attention has focused on the dangers of medium- and long-run imbalances in government budgets. This was less the case at the beginning of the recent financial crisis, when the question was if and how the government should try to stimulate the economy. But before long, the debate surfaced. Some argued that the government should prop up falling private demand with increased spending. Others claimed that increased government spending would have little to no stimulative effect in the short run and that it might even be contractionary.

Economists could offer little in the way of clarification, with venerated scholars falling on both sides of the debate. This failure of economists to agree on the issue leads some in the public to suppose that economists are incompetent, or perhaps worse, politically motivated.

The truth is that economists have struggled to answer the question, "What effect does an increase in government spending today have on output in the future?" In economics, this effect is called the government spending multiplier, and unfortunately for those of us who would like certainty on the matter, there are major challenges associated with measuring it. An appreciation for these challenges should explain why competent scholars can hold widely different opinions about the effect of government spending on output.

Measurement Challenges

Problems with measuring the government spending multiplier begin at the outset—with the way the question itself is phrased. At first, it seems like a natural question, but in fact it is far too general.

For starters, it is presumptuous to speak of "the" government spending multiplier as if there is only one. Because a change in government spending is likely to influence output over multiple periods in the future, separate multipliers could be created for each period. To calculate the appropriate multiplier, should we look at how much output changes one quarter in the future? One year? Five years? There is no universally accepted answer. Some studies report a collection of multipliers over a specific time period (for example, a multiplier for each quarter up to three years). Others average these numbers or report a range.

Taking a stand on timing is not sufficient however. One must also consider what type of government spending is increased. Surely an increase in military spending or spending on equipment will have a different effect on future output than an increase in infrastructure spending, education, or research. In practice, stimulus programs contain a mixture of spending types, so no two episodes are exactly the same. Most theoretical studies look at just the total level of spending and ignore these different uses.

A few have separated government spending into spending on consumption goods (like automobiles) and spending on investment goods (like infrastructure).

Our best estimates of the multiplier also depend upon a number of crucial assumptions about the environment in which the spending takes place. Is the economy in a recession? How is spending financed? How is contemporaneous monetary policy conducted? Are markets efficiently allocating resources, or is there room to improve the allocation? How are other countries responding? These are just a few of the important questions about context that affect the size of a multiplier.

Estimating the Multiplier Theoretically

The theoretical approach to estimating the government spending multiplier begins with a model. A model is a simplified representation of the economy designed to mimic aspects that are critical for answering a specific question. A good model includes as few variables as possible, but it must reproduce salient features of the data. A model that fits these criteria can be used like a laboratory to contemplate the circumstances under which government spending would boost GDP.

Models designed to address questions about the government spending multiplier have at least three fundamental components: utility-maximizing households, profit-maximizing firms, and a government. Households have an objective to maximize their utility over a given time horizon, subject to their lifetime budget constraint. That is, their goal is to obtain their most desired mix of goods over time while respecting their budget in every period. A typical budget would include wage income, income from savings, and transfers from the government.

Firms seek to maximize profit by employing workers and capital to create consumption goods, which households, the government, and, in some models, foreign consumers purchase. They also create new capital for use in future production.

The government's objective varies widely across models. Often it is left unspecified. In this case, the government mechanically raises tax revenues or issues new debt to cover its expenditures in each period. Typical expenditures are transfers to households, payments on outstanding debt, and spending on consumption or investment goods. Only this final activity constitutes "government spending," and often in models it is set by the researcher and is not determined by interactions of the other parts of the model. Typically, it is set to vary randomly around some trend.

A model containing the ingredients above will lead to the following basic macroeconomic identity, an expression which describes the relationship among the components.

> GDP Supplied = Consumption
> + Investment
> + Government Spending
> + Net Exports.

This identity is called the "aggregate resource constraint," and it simply states that all the production in an economy must be used somewhere. To put it succinctly, supply (the left-hand side) equals demand (the right-hand side). To better understand how government spending might affect GDP, it is helpful to consider how it would affect each component of the identity.

Demand

The components on the right-hand side of the aggregate resource constraint constitute the demand for goods and services in the economy within a period. Since the government is one of these components, increased government spending puts immediate upward pressure on demand for production; however, it has indirect effects on private activity as well, which show up in the other three components. Starting with consumption, new government spending will tend to push household demand for goods and services down. This is because additional government spending must be balanced by additional tax revenues either today or in the future. Either way, additional spending by the government means that, on average, households will have less disposable income over their lifetimes.

The reduction of consumption when a household's expected total lifetime income decreases is due to a phenomenon known as the "wealth effect." It is the economic force that causes someone to tighten their belt when they receive an unfavorable income surprise (in the case of a negative wealth effect; a positive effect would have the opposite impact). In the face of higher taxes, whether current or in the future, household consumption falls, and this dampens the upward pressure on aggregate demand caused by an increase in government spending.

The magnitude of the wealth effect will be governed by how much total lifetime income is expected to decline. In a benchmark scenario (the so-called Ricardian equivalence case), households fully offset new government spending by reducing consumption one-for-one. This leaves them with more savings, which they hold back to cover the future tax increases. This case of perfect substitution relies upon strong assumptions that are not true in the real world.

One assumption that would be necessary for Ricardian equivalence is that households live long enough to pay the future taxes that new government spending necessitates (or alternatively if parents care about their offspring exactly as if they were living forever). Studies that relax this assumption recognize that some households will not expect to pay the entire amount of future taxes. For these households, the wealth effect will be reduced.

The wealth effect can also be reduced if the tax burden is not evenly shared and households differ in their marginal propensity to consume. One recent study found evidence that the wealth effects of new government spending are larger for high-income households. These high-income households reduce their consumption sharply, but most

households do not. As a result, consumption does not decline very much in the aggregate.

New government spending is also thought to decrease the second component of aggregate demand, investment, for a couple of reasons. First, firms may anticipate an increase in business-related taxes. These new taxes would reduce the expected return from projects, suppressing firms' incentive to invest. Second, even if the new government spending is not financed with any business-related tax, the supply of savings available for investment will be partially reduced because some of it goes to government borrowing. In this way, government demand "crowds out" private sector investment. This crowding out of private domestic savings may be mitigated if additional savings flows in from foreign economies.

The final component, net exports, has traditionally been a small part of the US economy and so its effect on GDP has been given less attention. However, over time the trade sector has grown into a sizeable share of GDP, and the literature has given more thought to its role in shaping the size of the multiplier. The research suggests that, on net, one would anticipate a fall in the trade sector from new government spending. The wealth effect described above would reduce demand for imports. If interest rates rise however, foreign demand for dollars would cause the dollar to appreciate, making imports cheaper, and counteracting the wealth effect on imports to some degree. The expected change in imports then would be small. But because of the stronger dollar, exports would be expected to decline. Because there is no counteracting wealth effect on foreign consumers from domestic government spending, exports should unambiguously decline.

Overall, a rise in government spending would be expected to decrease aggregate consumption, aggregate investment, and net exports. The only factor increasing aggregate demand then is the direct effect of government spending. This is not the entire story, however. One must also consider the effect on the left-hand side of the aggregate resource constraint, the supply side.

Supply

In a standard macroeconomic model, production is a function of three inputs: technology, capital, and effective labor. An increase in any one of these inputs raises production.

Technology can be thought of as the knowledge that allows an economy to produce more given the same amount of capital and effective labor. However, increases in government spending don't much affect technology in the model. Generally, technology is set to increase steadily, with some small random accelerations and decelerations. Some models allow the rate of technological growth to be determined by firms' activities (for example, investment), but in these models technological growth is slow-developing, so temporary changes in government spending have little effect on it.

Likewise, capital is also thought of as slow to adjust. New machinery and buildings take time to construct; investment projects take time to plan and implement. For this reason, the capital stock within a period is usually modeled as fixed. The economy may develop new capital for the future by investing today, but within a given period, it must work with the capital available (from investment decisions made in the past).

Attention then must be turned to the third factor, effective labor. Effective labor is hours of work adjusted to account for differences in skill across the labor force (engineer hours are weighted more heavily than baristas'). Unlike the other two inputs, effective labor can be increased immediately. New workers can be hired, and current workers can be placed on overtime or part-time schedules or let go entirely. So would a rise in government spending increase production? As with most questions in economics, the answer is, "it depends."

The same wealth effect that causes households to cut back on consumption motivates them to consume less leisure as well. At the margin, households seek to pick up extra hours at work, get a second job, or send a working-age household member back into the labor force. When it comes to labor supply, however, there is a second force working in the opposite direction, resisting the wealth effect. Because the government must finance its increased spending, tax rates must rise to make up for the shortfall. This reduces the after-tax wage, which discourages additional work and encourages laborers to substitute leisure instead. The net effect depends on whether the wealth effect or the substitution effect dominates.

The strength of the substitution effect comes down to the timing of tax changes. If tax increases are pushed off into the future so that new debt finances current spending, then the current tax rate on wages will not change. This restrains the magnitude of the substitution effect, and effective hours are more likely to rise.

Equilibrium

Within the model framework, the overall effect on GDP of an increase in government spending comes down to the netting out of these multiple effects on the inputs. A rise in government spending is a direct increase on the demand side. This upward pressure however is dampened to some extent by decreased demand from other sectors. On the supply side, effective hours may increase or decrease.

In order for the aggregate resource constraint to be satisfied, all these forces must come into balance, meaning that something else in the model must adjust to bring everything into alignment. Typically, this is the model's interest rate, which may be nominal or real (adjusted for inflation), depending on the ingredients in the model. When the interest rate moves up, households save more and consume less. Firms, facing higher financing costs, hold off on new investment. A higher interest rate also encourages more labor since it provides more income for saving. Whether the interest rate increases or decreases in response to a government spending change varies from model to model.

An overview of the literature, like Ramey (2011), finds that a wide range of multipliers is possible depending upon how the model assumes spending is financed and how long spending remains above its average (figure 1). Unfortunately for policymakers, this range covers both positive and negative effects, meaning that it is unclear from theory whether an increase in government spending will lead to a rise in GDP.

Measuring the Multiplier Empirically

With theory returning an ambiguous answer about the size of the government spending multiplier, one must look to the data for answers. At first blush, it would seem very straightforward to measure the effect of government spending on output: Collect data on real GDP and on government spending; compare the two sets of data; and see if output increases above trend at or near the time government spending increased (figure 2). A number of challenges, however, make this exercise considerably more difficult.

In economics, and macroeconomics especially, we rarely, if ever, observe a true natural experiment. Unlike a chemist testing reactions in a lab, economists cannot experiment with government spending policies by shutting down particular sectors of the economy, introducing additional government spending, and recording the outcome. Instead, economists measure the effect of government spending on GDP through inference from past experience. Using mathematical techniques to filter out (or control for) other factors that might be influencing outcomes (such as monetary policy), economists try to identify which government spending changes are responsible for particular GDP changes.

But how does one know whether any particular movements in government spending and GDP are associated with each other? Just because we see both government spending and GDP rise, it does not mean that the former caused the latter. Perhaps an increase in GDP causes the government to spend more.

To work around this issue of causality, researchers focus on specific types of government spending that are thought to vary for reasons other than GDP changes. Military spending in particular is popular. There is little reason to think that conflicts overseas are closely linked to changes in US GDP. Because it is reasonable to believe that the causal relationship flows from military spending to GDP, it is possible to measure the effect on GDP from the rise in spending.

While focusing on military spending gets us past the causality problem, it creates to a new issue. There are many different ways a government can spend resources. Why should we think that every type of government activity has the same effect on GDP? How useful are results for military spending for estimating the effects of broad stimulus programs?

Finally, the issue of causality is also complicated by the possibility that instead of the increase in government spending causing changes in GDP, or the other way around, some third factor could be responsible for both. For instance, a persistent positive technology shock would boost GDP, which in turn would raise tax revenues and allow for increased government spending.

Figure 1. Researchers Have Gotten Different Results for the Multiplier

Number of studies

Value of the multiplier
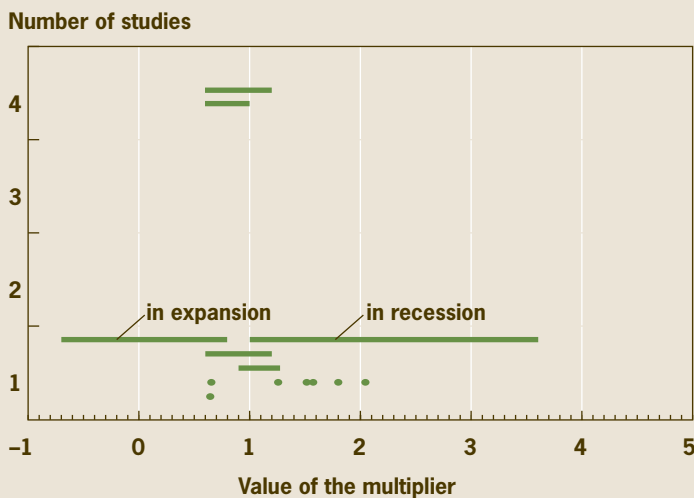
Source: Ramey, Valerie A., 2011. "Can Government Purchases Stimulate the Economy?" Journal of Economic Literature, 49(3): 673-85.

Figure 2. Changes in Government Spending Aren't Highly Correlated with Changes in GDP

Percent deviation from trend
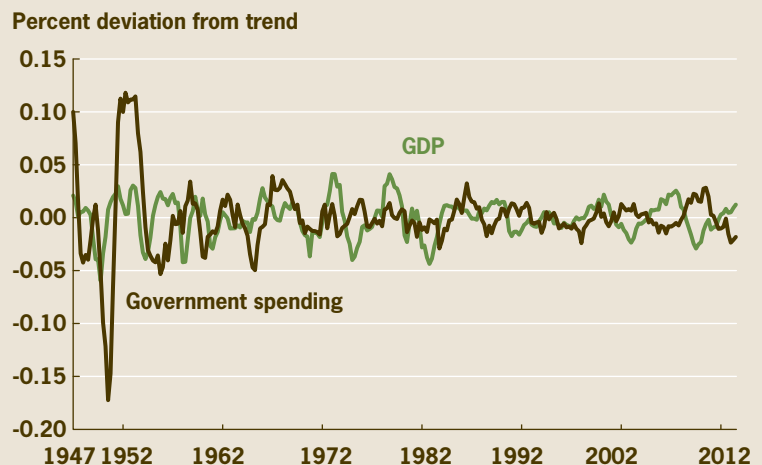
Source: Author's calculations using data from Bureau of Economic Analysis/Haver.

A second hurdle for identifying the effect of government spending on GDP empirically is public expectations. Often government spending packages are announced several quarters in advance, or they work their way through the legislative process for months. In either case, it is likely that by the time a program goes into effect and the money is actually spent, the public was well-informed long before. Households and firms, being forward-looking, may alter their current behavior in anticipation of future events. Firms could ramp up orders of raw materials and hire more workers to get ready for a big contract; and those newly hired workers may begin buying more goods with the new wage income.

If the public changes its behavior because it believes new government spending will be approved, and that change affects GDP over the time leading up to the program's implementation, an analyst may wrongly attribute those GDP movements to government spending at an earlier date. Moreover, because some of the impact from the program was smoothed out over the quarters leading up to the program taking effect, the GDP movement that is attributed to the government spending change is likely to be underestimated. (See figure 3.)

One way to approach these problems is to run a large vector autoregression (VAR). This is just a statistical process that looks for relationships (correlations) in the data. Given a lot of data on government spending, GDP, and a host of other potentially related variables (such as interest rates, measures of international trade, indicators of recession and expansion), the VAR assigns to each variable a number indicating how a small increase in that variable would change GDP based upon what has occurred in the past.
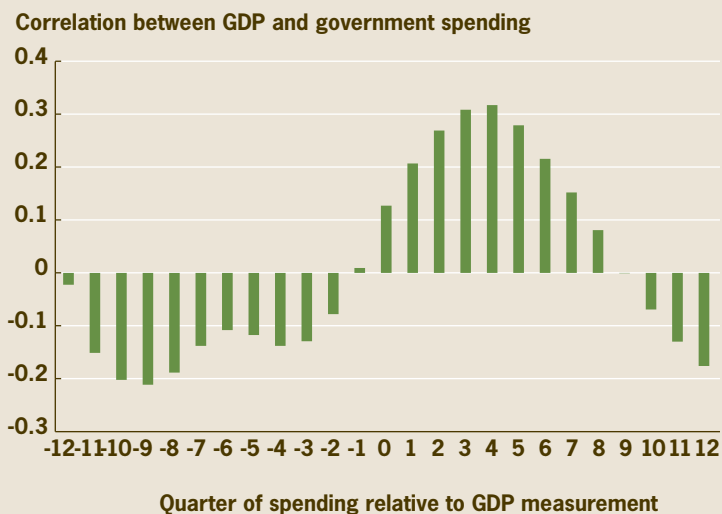
To resolve the expectations issue, the VAR approach makes some assumptions about how informed the public is about policy, and it includes restrictions on when government spending can affect GDP. Usually, government spending in a VAR is assumed to follow a process that is partly predetermined and partly random. The public expects the predetermined part, but the random component is a surprise, or in economist lingo, a "shock." Since shocks, by their nature, cannot be anticipated, identifying shocks and looking at how GDP responds to them, at least in theory, avoids the anticipation problem.

To address the timing problem, a VAR will include restrictions which reflect the researcher's beliefs about how government spending might affect GDP. For example, the VAR might be constructed so that only government spending shocks occurring in the previous four quarters can be assigned responsibility for changes in current GDP.

The ideal setup for a VAR is a very large data set taken from a similar environment and in which the variables display a lot of variation. As the number of observations or the variation gets smaller, it becomes harder for the VAR to precisely estimate the contribution of each variable. When the environment is considerably different from the present day (perhaps because of major changes to the role of government in the economy) then those data are less informative about what to expect from government spending today.

In the case of the United States, data are quite limited. There are only a few hundred quarterly observations, there is little variation in government spending after the Korean War, and the US economy has undergone some substantial changes over the last 70 years, which makes it hard to compare experiences across that time.

Because data limitations make it difficult to identify shocks, some economists turn to a more direct method for teasing the effect of government spending on GDP from the data called the "narrative approach" (Romer and Romer 1989). In this case, the researcher pores through congressional records on spending, news articles, or published market forecasts (Ramey 2011) to estimate public expectations. By comparing this expected path for government spending to the actual path, the researcher can construct a sequence of "shocks," and then examine how GDP responds to them.

Ultimately, all the issues with the empirical approach boil down to what statisticians call identification. To avoid the anticipation problem, an economists needs to identify government spending shocks. To address the timing problem, an economist needs to assign the correct shocks to the correct GDP changes. Any resolution requires some subjective judgment on the part of the economist, which immediately opens any results to debate. In addition, even supposing that a consensus approach did exist, it would almost certainly require an economist to discard a lot of the data. Limited data allow for less precise statistical conclusions. Across studies, the range of estimates of the government spending multiplier is wide; and within studies, the range of statistically plausible values is often wide as well.

Figure 3. Even at Various Leads and Lags, Government Spending and GDP Aren't Highly Correlated

Correlation between GDP and government spending

Quarter of spending relative to GDP measurement

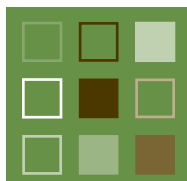Source: Author's calculations using data from Bureau of Economic Analysis/Haver.

Next Steps

Economists have made considerable progress toward categorizing the situations in which the government spending multiplier is greater than 1 and understanding the mechanisms that would produce the result. From the view of theory, in order for government spending multipliers to be high, the wealth effect needs to be large since a large wealth effect increases hours worked (or labor input) in the near term. The longer government spending remains away from its trend or the more evenly the costs of new spending fall across households, the more likely this condition is to be satisfied. Differences in modelling assumptions across the literature lead to differing conclusions about the government spending multiplier.

The challenge facing the profession is to empirically distinguish between these models. Better measurements of how GDP and government spending move together could lead to some models being rejected and possibly shift the preponderance of evidence in favor of one type of model over another. As yet, empirical hurdles leave plenty of room for debate over the size of the multiplier.

References

Anderson, Emily, Atsushi Inoue, and Barbara Rossi, 2013. "Heterogeneous Consumers and Fiscal Policy Shocks," manuscript, Duke University, NCSU, and UPF.

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

Christina D. Romer and David H. Romer, 1989. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz." *NBER Macroeconomics Annual,* vol. 4, pp. 121-170.